

QUADERNI DELLA RIVISTA

RIVISTA DELLA CORTE DEI CONTI



CORTE DEI CONTI



Intelligenza artificiale

Impiego e implicazioni
sotto il profilo tecnologico, etico e giuridico



Quaderno n. 2/2024

Direttore responsabile: Tommaso Miele

Direttore scientifico: Francesco Saverio Marini

Comitato scientifico

Nicolò Abriani – Federico Alvino – Luca Antonini – Miguel Azpitarte Sánchez – Luigi Balestra – Daniela Bolognino – Francesco Capalbo – Marcello Cecchetti – Stefano Cerrato – Vincenzo Cerulli Irelli – Fabio Cintioli – Giacomo D’Attorre – Alberto Dello Strologo – Ruggiero Dipace – Gianluca Maria Esposito – Gabriele Fava – Francesco Fimmanò – Enrico Gabrielli – Franco Gallo – Edoardo Giardino – Francesco Grieco – Giovanni Guzzetta – Margherita Interlandi – Amedeo Lepore – Alberto Lucarelli – Massimo Luciani – Francesco Saverio Marini – Bernardo Giorgio Mattarella – Giuseppe Melis – Donatella Morana – Nino Paolantonio – Fulvio Pastore – Antonio Pedone – Giovanni Pitruzzella – Aristide Police – Stefano Pozzoli – Nicola Rascio – Giuseppe Recinto – Guido Rivosecchi – Aldo Sandulli – Maria Alessandra Sandulli – Franco Gaetano Scoca – Gennaro Terracciano – Raffaele Trequattrini – Antonio Felice Uricchio – Nicolò Zanon

Quaderni della Rivista

Coordinatore: Alberto Avoli

Redazione: Ernesto Capasso

Editing

Anna Rita Bracci Cambini

Agnese Colelli – Stefano De Filippis – Valeria Gallo – Lucia Pascucci

La Rivista della Corte dei conti è a cura del Servizio Massimario e Rivista

La rivista è consultabile anche in:
www.rivistacorteconti.it

RIVISTA
DELLA
CORTE DEI CONTI
QUADERNI

Quaderno n. 2/2024

Direttore responsabile
Pres. Tommaso Miele

SOMMARIO

- Guido Carlino, <i>Introduzione</i>	3
- Ginevra Ferrina Ceroni, <i>La dignità della persona: tra Gpdr e AI act</i>	5
- Teresa Numerico, <i>Per una genealogia critica dell'intelligenza artificiale</i>	13
- Lucia de Grimani, <i>Ai e futuro occupazionale: sfide e opportunità in un mondo in evoluzione</i>	23
- Salvatore Gaglio, <i>Intelligenza artificiale. Una sintesi aggiornata</i>	29
- Oreste Pollicino, <i>Disinformazione e intelligenza artificiale: un cocktail esplosivo?</i>	41
- Elena Tomassini, <i>Intelligenza artificiale, esimente politica e programmazione della spesa</i>	51
- Daniela Anselmi, Federico Smerchinich, <i>L'algoritmo, tra intelligenza artificiale e diritto amministrativo</i>	55
- Vito Tenore, <i>Una riflessione sistematica: l'intelligenza artificiale può sostituire un giudice o trasformarlo da homo sapiens in homo numericus?</i>	61
- Maria Teresa D'Urso, <i>Il d.d.l. (n. 1066 A.S.), di iniziativa governativa, approvato il 23 aprile 2024, sulla intelligenza artificiale. i principi fondamentali dell'Ai per il suo utilizzo in Italia, in linea con "l'Ai act" deliberato dall'Ue. Le disposizioni di settore. la strategia nazionale e le nuove agenzie istituite. la tutela per gli utenti e per il diritto d'autore e le pene per i trasgressori. il G7 svoltosi il 13-15 giugno 2024 a Borgo Egnazia (BR), con l'intervento di Papa Bergoglio</i>	97
- Pasquale Fava, <i>La c.d. "intelligenza artificiale": personalità elettronica, imputazioni e responsabilità tra diritto civile, penale ed amministrativo</i>	109
- Antonello Lirosi, <i>L'intelligenza artificiale nel diritto amministrativo – tra riserva di umanità e necessità di garantire una maggiore efficienza amministrativa</i>	127
- Paolo Evangelista, <i>Intelligenza artificiale e responsabilità amministrativo-contabile</i>	139
- Pierfrancesco Miele, <i>Intelligenza artificiale e pubblica amministrazione</i>	143
- Giovanna Capilli, <i>Intelligenza artificiale e ricerca del precedente: un aiuto a transigere?</i>	161
- Alberto Avoli, <i>Conclusioni</i>	167

INTRODUZIONE

GUIDO CARLINO

Presidente della Corte dei conti

Nel presente quaderno sono affrontati, sotto varie angolazioni e con apprezzabile chiarezza espositiva, diversi profili riguardanti un fenomeno scientifico tra i più rilevanti e complessi del nostro tempo, l'impiego dell'intelligenza artificiale (IA) e le implicazioni che ne derivano sotto il profilo tecnologico, etico e giuridico.

I numerosi contributi qui raccolti offrono riflessioni di stimati autori che esaminano le reazioni alle sfide poste dall'utilizzo dell'intelligenza artificiale. Il risultato è una proficua sinergia tra diverse discipline che interagiscono e procedono verso nuove frontiere digitali, dal diritto amministrativo, che è direttamente attraversato dagli effetti giuridici del crescente dilagare delle tecnologie, a quello costituzionale, che chiama in causa i diritti fondamentali, alle scienze ingegneristiche.

La rivoluzione digitale apre un inedito e affascinante capitolo con ricadute impressionanti in tutti i campi della vita quotidiana delle persone, influenzandone sotto diversi aspetti la qualità.

Trasformando profondamente il modo di lavorare, la digitalizzazione si è rivelata un formidabile strumento in grado di potenziare le diverse abilità, con uno spettro di applicazioni possibili assai ampio che permea tutti i settori, compresa l'amministrazione pubblica, migliorandone l'efficienza, la trasparenza, la veridicità e l'imparzialità.

L'adozione di procedure amministrative telematiche uniformi sul territorio, anche con funzionali modalità da remoto, facilita ai cittadini e alle imprese la fruizione più tempestiva dei servizi, riducendo sprechi e costi, a fronte di una maggiore qualità e apre la strada a un rinnovato rapporto tra amministrazioni pubbliche, operatori e cittadini, ai quali poter fornire una risposta più rapida e adeguata.

Una leva strategica, dunque, intesa a migliorare radicalmente l'azione amministrativa, razionalizzando la programmazione delle scelte nei processi decisionali.

L'utilizzo dell'intelligenza artificiale, infatti, mira a selezionare la soluzione più probabile, cioè credibile o verosimile, in base a informazioni pregresse e sicure che la presentano come la migliore possibile; consente, dunque, di volgere lo sguardo al futuro attraverso il passato.

È, in pratica, in atto un passaggio dalla cittadinanza tradizionale a quella digitale e tecnologica, che porta a dover ridisegnare nuove relazioni tra cittadini, imprese e istituzioni.

Va, tuttavia, considerata l'estrema delicatezza dei profili connessi all'applicazione di algoritmi all'azione pubblica. In tale ambito, le Istituzioni hanno la responsabilità di garantire che il processo decisionale che si avvale dell'intelligenza artificiale non sia lasciato indipendente e venga controllato in ogni sua fase, restando ancillare e servente rispetto alla valutazione, che deve essere sempre appannaggio dell'operatore umano e rivolta alla soddisfazione del bene comune, senza mai prescindere dai valori etici e da solide catene valoriali.

In ogni caso va preservato il rispetto dell'equilibrio tra macchina e persona, tra efficienza amministrativa e salvaguardia della riserva di umanità, nell'osservanza dei principi di uguaglianza e di non discriminazione, imperativi caratteristici del decisore pubblico.

Ciò, al fine di assicurare, altresì, adeguati livelli di tutela agli interessi pubblici e al cittadino mediante un utilizzo consapevole e responsabile dei dati personali, che vanno protetti a salvaguardia imprescindibile della dignità umana e dei diritti fondamentali.

La rivoluzione digitale procede a una velocità così incalzante che il diritto fatica a tenere il passo e diviene più che mai necessario ben delineare le regole che governano l'utilizzo dei sistemi artificiali.

L'intelligenza artificiale postula in particolare un'amministrazione sempre più attenta e consapevole della complessità della realtà e dei fattori sociali, scientifici, economici e tecnici coinvolti.

La disponibilità di numerose banche dati, l'impiego di algoritmi predittivi e l'adozione di standard di alta qualità basati sull'utilizzo di piattaforme interoperabili richiedono agenti pubblici sempre più preparati e diligenti.

È necessario, pertanto, assicurare un'interazione tra l'informatica e l'intelligenza umana, posto che, nell'ambito di scelte logiche elaborate o elaborabili sulla base delle risultanze dell'intelligenza artificiale, compete all'agente pubblico la selezione di quella più coerente con l'interesse pubblico.

Non va infatti trascurato che uno dei problemi più concreti che sorgono dalle scelte fondate sull'intelligenza artificiale è che non sempre esse forniscono la motivazione che la legge vuole alla base di ogni scelta e provvedimento amministrativo.

È per tale motivo che ogni sistema di intelligenza artificiale non può essere lasciato indipendente nel suo agire, ma ogni output prodotto deve essere sempre controllato dall'operatore umano che possa correggerne l'errore, utilizzando la propria intelligenza ed esperienza.

Proprio questa interazione, che esalta il ruolo dell'uomo, fonda il presupposto anche per l'eventuale insorgenza della responsabilità amministrativa per colpa grave nel caso di erroneo o improprio utilizzo dell'uso di tecnologie, programmi ed algoritmi che, per quanto si possano rivelare utili, non possono fare passare in secondo piano l'importanza e la centralità della decisione umana.

Non solo l'attività amministrativa ma anche quella giurisdizionale beneficia dei molteplici vantaggi recati dall'intelligenza artificiale. È, infatti, possibile ridurre la durata dei procedimenti giudiziari, agevolare la ricerca giurisprudenziale o dottrinale e muovere i primi cauti passi verso la predittività degli orientamenti della giurisprudenza (in ciò accrescendo il ruolo della nomofilachia), fermo restando che alla valutazione degli atti processuali debba sempre provvedere il giudice utilizzando il proprio bagaglio culturale e valoriale, potendo, al più, l'intelligenza artificiale costituire un mero ausilio tecnologico.

La questione dell'applicazione dell'intelligenza artificiale nel settore della giustizia ha dato luogo all'emanazione, avvenuta il 4 Dicembre 2018, da parte della Commissione Europea per l'Efficacia della Giustizia (CEPEJ) del Consiglio d'Europa, della "Carta Etica Europea per l'uso dell'intelligenza artificiale nei sistemi di giustizia penale e nei relativi ambienti", che ha evidenziato la necessità che siano chiare e sindacabili le regole in base alle quali il processo è stato impostato, dalla costruzione all'inserimento dei dati, alla loro validità e gestione.

Ben consapevole delle utilità recate dall'innovazione tecnologica, la Corte dei conti ha avviato da tempo un processo di trasformazione digitale, sia nell'ambito delle proprie funzioni istituzionali di controllo e giurisdizionali, sia all'interno delle linee strategiche dell'azione amministrativa, unitamente a studi per avvalersi dei nuovi strumenti offerti dall'intelligenza artificiale.

La raccolta dei dati giudiziari attraverso strumenti tecnologici sta andando, infatti, verso una radicale trasformazione delle modalità con cui la giurisdizione acquisisce, ad esempio, la conoscenza dei fatti rilevanti, oltre che verso l'ampliamento della capacità di interloquire con l'autorità e con i controinteressati.

In proposito, è in via di sperimentazione l'anonimizzazione automatizzata dei dati personali nelle sentenze, al fine di proteggere la *privacy* e la riservatezza degli individui coinvolti nei procedimenti giudiziari.

Dall'ausilio dei modelli predittivi potrà trarre giovamento anche la funzione di controllo, messa in grado di individuare situazioni di criticità finanziaria potenzialmente idonee a determinare squilibri finanziari ovvero il dissesto dell'ente locale.

Anche sotto il profilo degli accertamenti svolti dal requirente contabile, l'agilità tecnologica, nel definire in maniera chiara e lineare i percorsi procedurali, consente anche una più agevole identificazione delle tecniche di imputazione degli atti e delle correlate responsabilità.

Concludendo, va rilevato che il consolidamento della rivoluzione digitale non può prescindere dal superamento delle connotazioni tecnicistiche dei processi informatici, dalle quali deriva una autoreferenzialità, che finisce per escludere larghi strati della popolazione. Le innovazioni, infatti, modificano realmente un sistema quando coinvolgono e sono adottate dalla maggior parte della comunità.

L'appropriata ed etica applicazione di procedure automatizzate e di tecnologie avanzate può rivelarsi uno strumento di amplificazione delle libertà e dei diritti, compresi quelli fondamentali di ogni individuo, dei quali va assicurata l'effettività.

In attesa della dimensione istituzionale che verrà acquisita dal fenomeno digitale, il mondo del diritto comincia a lasciare emergere alcune delle necessarie chiavi di governo.

Già gli stessi Padri costituenti vollero incoraggiare l'uso della tecnologia, statuendo all'art. 9 della Costituzione che "*La Repubblica promuove lo sviluppo della cultura e della ricerca scientifica e tecnica*", a tutela del progresso e, dunque, della qualità della vita.

La normativa primaria si è in parte adeguata, e sempre di più sarà tenuta a farlo con il fenomeno al nostro esame, al dettato costituzionale e all'evoluzione scientifica e sociale.

Con solidi referenti normativi si possono ridurre incertezze e perplessità in materia, che inducono la giurisprudenza ad applicare principi già esistenti a un fenomeno completamente nuovo, con sicure ricadute anche sui contenuti tradizionali degli istituti di riferimento.

È, pertanto, fondamentale che la progressiva adozione di modelli tecnologici di intelligenza artificiale cammini di pari passo con una regolazione certa e definita, via via emanata a livello globale, che assicuri conoscenze, formazione, controllo e responsabilità, a garanzia di benefici equamente distribuiti.

* * *

LA DIGNITÀ DELLA PERSONA: TRA GPDR E AI ACT

di Ginevra Ferrina Ceroni (*)

Abstract: La dignità personale è al centro del costituzionalismo europeo e della regolazione del fenomeno digitale, rappresentando un filo conduttore sia nel Gdpr che nell'AI Act. Questi strumenti normativi riflettono l'importanza della protezione dei dati personali, integrata nei Trattati europei, nella tutela della dignità umana. Il presente contributo analizza i principi e gli istituti principali del Gdpr e dell'AI Act che proteggono la dignità personale, oltre agli orientamenti giurisprudenziali dell'Autorità Garante per la protezione dei dati personali in vari settori, pubblici e privati, dove emerge l'intersezione tra garanzie algoritmiche e difesa della dignità umana. Vengono inoltre esaminati i profili critici della governance dell'IA nell'AI Act, sottolineando che l'efficacia della tutela della dignità personale dipende dal controllo e dalla vigilanza delle Autorità competenti. Si auspica la designazione delle Autorità di protezione dati personali come Autorità competenti per la vigilanza sui sistemi di IA ad alto rischio, usati per attività di contrasto, gestione delle frontiere, giustizia e democrazia, in conformità con l'art. 74, par. 8, e come "autorità garante dei diritti fondamentali" ai sensi dell'art. 77 dell'AI Act.

Personal dignity lies at the heart of European constitutionalism and the regulation of the digital phenomenon, serving as a common thread in both the Gdpr and the AI Act. These regulatory frameworks underscore the importance of personal data protection, enshrined in European Treaties, for safeguarding human dignity. This contribution examines the key principles and provisions of the Gdpr and the AI Act that uphold personal dignity, alongside jurisprudential guidelines from the Data Protection Authority across various public and private sectors, highlighting the intersection between algorithmic safeguards and the protection of human dignity. Additionally, it scrutinizes critical aspects of AI governance within the AI Act, emphasizing that the effectiveness of safeguarding personal dignity hinges on oversight and vigilance by designated competent authorities. It advocates for designating data protection authorities as competent entities to oversee high-risk AI systems used in law enforcement, border management, justice, and democracy, in accordance with Article 74(8), and as national public authorities protecting fundamental rights under Article 77 of the AI Act

Sommario: 1. Cenni introduttivi: la dignità della persona al centro della contemporanea regolazione digitale. – 2. La declinazione della dignità umana nella sistematica del Gdpr e gli interventi dell'Autorità Garante in materia di protezione dei dati personali. – 3. La tutela della dignità personale all'interno dell'AI Act, tra prassi di IA vietate e "ad alto rischio". – 4. La sfida della governance in materia di IA Conclusioni e prospettive future.

1. Cenni introduttivi: la dignità della persona al centro della contemporanea regolazione digitale

La dignità della persona occupa uno spazio centrale nella storia millenaria della filosofia morale, della teologia, del diritto. È un archetipo, in quanto da essa originano tutti i diritti. Si tratta di un concetto aperto, dinamico, metagiuridico e insuscettibile di definizione (1): un meta-principio sul quale si fondano diverse Carte costituzionali e internazionali (2).

(*) G. Ferrina Ceroni è vicepresidente dell'Autorità garante per la protezione dei dati personali; professore ordinario di diritto costituzionale italiano e comparato presso l'Università di Firenze.

(1) Cfr. C. Mortati, *Commento all'art. 1*, in *Comm. Cost. Branca*, I, Bologna-Roma, 1975; C. Mortati, *Istituzioni di diritto pubblico*, I, Padova, 1975. Cfr. altresì A. Ruggeri, A. Spadaro, *Dignità dell'uomo e giurisprudenza costituzionale (prime notazioni)*, in *Pol. dir.*, 1991, 347; M. Bellocci, P. Passaglia (a cura di), *La dignità dell'uomo quale principio costituzionale*. Quaderno predisposto in occasione dell'incontro trilaterale delle Corti costituzionali italiana, spagnola e portoghese, tenutosi a Roma il 30 settembre e l'1 ottobre 2007; E. Ceccherini (a cura di), *La tutela della dignità dell'uomo. Lezioni Volterranne* 2006, Napoli, 2008; A. Argiroffi, P. Becchi, D. Anselmo, (a cura di), *Colloqui sulla dignità umana*, Roma, 2008; A. Pirozzoli, *Il valore costituzionale della dignità. Un'introduzione*, Roma, 2007; U. Vincenti, *Diritti e dignità umana*, Roma-Bari, 2009; Q. Camerlengo, *Contributo ad una teoria del diritto costituzionale cosmopolitico*, Milano, 2007, 338 ss.; U. Vincenti, *Diritti e dignità umana*, Roma-Bari, 2009. Cfr. da ultimo A. Occhipinti, *Tutela della vita e dignità umana*, Torino, 2008, 34, secondo cui il concetto di dignità non è fisso, ma si adegua al mutare dei tempi e sintetizza: "il livello di sensibilità espressa dalla società ed il senso del rispetto dovuto alla persona, ricostruito in relazione alla fase temporale di riferimento".

(2) Se lo si colloca in una prospettiva storica, il valore della dignità umana emerge prepotentemente a livello costituzionale a partire dalla seconda metà del '900, come diretta reazione sia agli orrori della Shoah, sia ai fascismi e ai totalitarismi che hanno attraversato il continente europeo. Non è casuale che si ritrovino riferimenti costituzionali alla dignità personale all'art. 1 della Cost. tedesca, il quale sancisce che: "La dignità dell'uomo è intangibile". Sul concetto di "intangibilità" della dignità nella Cost. tedesca, cfr. F. Berardo, *La dignità umana è intangibile: il dibattito costituente sull'art. 1 del Grundgesetz*, in

All'interno della Costituzione italiana manca una specifica norma che tuteli in quanto tale la dignità umana. Tale concetto è tuttavia richiamato in varie disposizioni (artt. 32, 36, 41 Cost.) e la sua protezione costituisce il corollario dell'idea fondante della Repubblica. Inoltre, la dignità sociale costituisce la pietra angolare su cui viene a fondarsi il principio di uguaglianza dell'art. 3 Cost.; ovviamente è il fondamento sostanziale dell'art. 2 Cost. e presupposto implicito di tutti i diritti di matrice costituzionale. Tanto che in una nota sentenza della Corte costituzionale (3) si afferma che la dignità umana è un valore costituzionale che permea di sé tutto il diritto positivo.

Si può senz'altro affermare che la dignità umana costituisce il cuore del principio personalista del costituzionalismo contemporaneo da intendersi, da un lato, come centralità della persona umana nell'impianto costituzionale. Dall'altro, come individualismo giuridico, ossia l'individuo è il fondamento di tutto l'impianto statale di azione e di tutela giuridica.

A livello internazionale, il Preambolo della Carta delle Nazioni unite riafferma la fede dei popoli nella "dignità e nel valore della personalità umana" (4). Il principio della dignità della persona umana riecheggia anche nel Preambolo e nell'articolo 1 della Dichiarazione universale dei diritti umani (5) e nella Convenzione europea dei diritti dell'uomo (Cedu).

Nell'ambito dell'Unione europea, la dignità personale permea il dettato della Carta dei Diritti Fondamentali dell'Unione europea. Nel Preambolo di quest'ultima viene affermato che l'Unione si fonda sui valori indivisibili di dignità umana, di libertà, di uguaglianza e di solidarietà (6). La Carta muove quindi dal presupposto fondamentale dell'immanenza della dignità umana; perciò, l'Unione europea pone la persona al centro della sua azione regolatoria.

Non deve, dunque, stupire che il legislatore europeo abbia dedicato una particolare attenzione all'intangibilità della dignità personale, ponendo tale valore al centro di due strumenti normativi cruciali nella contemporanea regolazione del fenomeno digitale, ossia il Regolamento europeo per la protezione dei dati personali ("Gdpr") (7) e il Regolamento europeo sull'intelligenza artificiale ("AI Act") (8). La centralità della tutela della dignità personale nei due strumenti si estrinseca nella protezione dei dati personali, divenuta parte integrante dei Trattati europei. Infatti, sia il Gdpr che l'AI Act si fondano sugli artt. 8 Cdfue e 16 Tfue, nella misura in cui la protezione dei dati personali assurge a baluardo imprescindibile di garanzia della dignità personale, orientando l'azione regolatoria in materia tecnologica.

In un mondo sempre più datificato, in cui le nostre identità sono "fatte di dati", la protezione dei dati personali diventa, dunque, strumentale alla difesa irrinunciabile della dignità personale.

L'obiettivo perseguito dal legislatore europeo sia con il Gdpr che con l'AI Act è quello – estremamente ambizioso – di trovare un punto di equilibrio tra protezione della dignità personale e sostegno all'innovazione tecnologica. Mantenere l'uomo al centro dello sviluppo tecnologico, promuovere l'umanesimo digitale e incoraggiare uno sviluppo antropocentrico dell'IA: lungi dal costituire vuote formule di stile, costituiscono concetti e valori centrali, seppur con le dovute differenze, nella sistematica tanto del Gdpr, quanto dell'AI Act.

Quad. cost., 2, 2006, 387-400. Si veda anche la Costituzione del Portogallo, che all'art. 1 stabilisce che "il Portogallo è una Repubblica sovrana fondata sui principi della dignità umana e della volontà del popolo", oppure l'art. 10, c. 1, della Costituzione spagnola, secondo cui: "la dignità della persona, i diritti inviolabili ad essa inerenti, il libero sviluppo della personalità, il rispetto della legge e dei diritti altrui sono il fondamento dell'ordine politico e della pace sociale".

(3) Corte cost., 11 luglio 2000, n. 293.

(4) Preambolo della Carta delle Nazioni Unite, 1945: "Noi popoli delle Nazioni Unite, decisi a salvare le future generazioni dal flagello della guerra, che per due volte nel corso di questa generazione ha portato indicibili afflizioni all'umanità, a riaffermare la fede nei diritti fondamentali dell'uomo, nella dignità e nel valore della persona umana, [...]".

(5) Art. 1, Dichiarazione universale dei diritti umani, 1948: "Tutti gli esseri umani nascono liberi ed eguali in dignità e diritti. Essi sono dotati di ragione e di coscienza e devono agire gli uni verso gli altri in spirito di fratellanza".

(6) Preambolo della Carta dei diritti fondamentali dell'Unione europea, 2012: "Consapevole del suo patrimonio spirituale e morale, l'Unione si fonda sui valori indivisibili e universali della dignità umana, della libertà, dell'uguaglianza e della solidarietà; essa si basa sul principio della democrazia e sul principio dello Stato di diritto. Pone la persona al centro della sua azione istituendo la cittadinanza dell'Unione e creando uno spazio di libertà, sicurezza e giustizia". Con riferimento alla Carta dei diritti fondamentali, si vedano le riflessioni di F. Sacco, *Note sulla dignità umana nel "diritto costituzionale europeo"*, in S. Panunzio (a cura di), *I diritti fondamentali e le Corti in Europa*, Napoli, 2005, 596 ss. Cfr. anche M. Olivetti, *Art. 1. Dignità umana*, in R. Bifulco, M. Cartabia, A. Celotto (a cura di), *L'Europa dei diritti. Commento alla Carta dei diritti fondamentali dell'Unione Europea*, Bologna, 2001, 45; R. Bifulco, *Dignità umana e integrità genetica nella Carta dei diritti fondamentali dell'Unione Europea*, in *Bioetica*, 3, 2003, 443-478.

(7) Reg. Ue 27 aprile 2016, n. 679, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati personali).

(8) Reg. Ue 13 giugno 2024, n. 1689, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i reg. (Ce) n. 300/2008, (Ue) n. 167/2013, (Ue) n. 168/2013, (Ue) 2018/858, (Ue) 2018/1139 e (Ue) 2019/2144 e le direttive 2014/90/Ue, 2016/797 e 2020/1828 (regolamento sull'intelligenza artificiale).

2. La declinazione della dignità umana nella sistematica del Gdpr e gli interventi dell'Autorità Garante in materia di protezione dei dati personali

Nella sistematica del Gdpr, la tutela della dignità personale può essere rintracciata nel diritto fondamentale all'autodeterminazione informativa dell'individuo, sancito per la prima volta in un'originale sentenza della Corte costituzionale tedesca (9). All'interno del Gdpr, l'autodeterminazione informativa si declina concretamente nell'enucleazione di una serie di principi che devono permeare i trattamenti e nell'attribuzione al soggetto interessato dal trattamento di un novero di diritti azionabili in relazione ai propri dati personali.

Come noto, l'art. 5 del Gdpr (10) stabilisce la liceità, la correttezza e la trasparenza dei trattamenti, la limitazione delle finalità e della conservazione dei dati personali, la minimizzazione, l'esattezza, l'integrità e la responsabilizzazione (c.d. *accountability*) del titolare. Si tratta di principi fondamentali posti a garanzia degli interessati, già previsti dalla Direttiva 95/46/Ce (11) e sanciti a livello di Consiglio d'Europa dalla Convenzione 108 (12) con riferimento all'elaborazione automatizzata delle informazioni. Tali principi devono essere osservati dai titolari del trattamento mediante l'attuazione di concrete misure tecniche ed organizzative e l'adempimento di specifici obblighi, espressione e corollari dei suddetti principi. Si pensi, a titolo esemplificativo, al principio di trasparenza previsto dall'art. 5, par. 1 lett. a) del Gdpr, che si traduce nell'obbligo di rendere informativa agli interessati per renderli consapevoli della circostanza che sia in corso un trattamento di dati che li riguarda e di quali diritti possono esercitare. O ancora, la minimizzazione impone ai titolari di effettuare delle scelte operative volte a ridurre, per quanto possibile, la raccolta e l'uso di dati personali, prevenendo prassi di trattamenti a strascico o massivi di dati. Infine, si considerino i principi di integrità e riservatezza, i quali da astratti valori si traducono in misure concrete, quali installare un anti-malware, mantenere copie di *back-up* dei dati, chiudere un armadio a chiave e così via.

È inoltre attribuito all'interessato un ventaglio di diritti esercitabili in relazione ai propri dati personali come l'accesso, l'oblio, la rettifica, l'opposizione, la limitazione e la portabilità (13). Assumendo una connotazione procedurale, tali diritti attuano in concreto l'autodeterminazione informativa e sono funzionali all'esercizio di altri diritti e libertà fondamentali di rango costituzionale. Si pensi al caso del paziente che desidera accedere e richiedere una copia della propria documentazione medica, in cui l'esercizio del diritto di accesso sia strumentale alla protezione del diritto alla salute. Oppure si consideri l'oblio e la rettifica dei dati personali che sono funzionali a proteggere l'identità e la reputazione della persona fisica lesa da una notizia pubblicata online, divenuta datata con lo scorrere del tempo (14).

Tra i diritti attribuiti dal Gdpr agli interessati, uno tra tutti merita particolare attenzione in relazione alle prassi sempre più invasive dei trattamenti automatizzati, della profilazione e degli usi dell'IA: il diritto di accesso, spiegazione e contestazione delle logiche automatizzate e algoritmiche previsto dall'art. 22 Gdpr (15).

Il connubio indissolubile tra garanzie algoritmiche e difesa della dignità personale è evidente se si ragiona in termini di possibile pregiudizio agli interessati derivante dall'assunzione di decisioni automatiche e se si guarda agli indirizzi giurisprudenziali dell'Autorità Garante. Facendo applicazione del previgente art. 14 del d.lgs. n. 196/2003 (antecedente alla novella legislativa apportata con d.lgs. n. 101/2018) e poi dell'art. 22 del Gdpr, fin dal 2013 l'Autorità Garante ha applicato estensivamente e coerentemente le disposizioni previste in materia di trattamenti automatizzati, anticipando in diversi settori principi e tutele che sarebbero confluiti nell'*AI Act*. Di seguito si fa riferimento ad alcuni interventi del Garante particolarmente significativi in tal senso.

(9) Corte cost. tedesca, sentenza 15 dicembre 1983. Riconosciuta come “una straordinaria dimostrazione di attivismo giuridiale”, la Corte dichiarava la parziale incostituzionalità di un censimento della popolazione tedesca pianificato dal 1981, riconoscendo l'esistenza di un nuovo diritto fondamentale, quello alla “autodeterminazione informativa” (“*Recht auf informationelle Selbstbestimmung*”). La pronuncia ha avuto un'ampia eco nel dibattito giurisprudenziale e dottrinale tedesco ed internazionale in tema di diritti della personalità.

(10) Cfr. art. 5, par. 1 e 2, Gdpr.

(11) Cfr. art. 6, par. 1 e 2, direttiva 95/46/Ce.

(12) Cfr. art. 5, par. 1, Convenzione 108.

(13) Cfr. artt. 15-22 Gdpr.

(14) La giurisprudenza dell'Autorità Garante in materia di diritto all'oblio è cospicua. Senza pretesa di esaustività, si vedano tra i più recenti provvedimenti: n. 297 del 9 maggio 2024 (doc. web n. 10027850); n. 275 del 9 maggio 2024 (doc. web n. 10027521); n. 274 del 9 maggio 2024 (doc. web n. 10022403); n. 109 del 23 marzo 2023 (doc. web n. 9896373); n. 146 del 7 marzo 2024 (doc. web n. 10009352); n. 64 del 2 marzo 2023 (doc. web n. 9874462).

(15) Cfr. art. 22 Gdpr. Si vedano altresì i cons. 4, 24, 38, 60, 63, 71, 72, 73, 91. Cfr. inoltre Gruppo di Lavoro Articolo 29, Linee guida in materia di processi decisionali automatizzati e profilazione, adottate il 3 ottobre 2017. La *ratio* della disposizione è analoga a quella che mosse il legislatore europeo ad adottare l'art. 15 della previgente direttiva 95/47/Ce, ossia proteggere l'individuo nell'ambito di processi decisionali che lo riguardano unicamente fondati sulla base della sua “data shadow”. Per un commento autorevole sull'art. 22 del Gdpr, cfr. L. A. Bygrave, *Article 22. Automated individual decision-making including profiling*, in C. Kuner, L. A. Bygrave, C. Docksey, *The EU General Data Protection Regulation (Gdpr): A Commentary*, Oxford, 2019.

Nel 2013 il Garante si è occupato del caso del c.d. “Redditometro”, uno strumento di controllo volto a combattere l’elusione fiscale (16). L’algoritmo impiegato dall’Agenzia delle Entrate selezionava i contribuenti da sottoporre ad accertamento, per rideterminarne il reddito effettivo sull’imputazione al contribuente di spese presunte, quantificate sulla base dell’attribuzione di un profilo (*cluster*) ricavato anche dalla sua appartenenza ad una specifica tipologia di famiglia e dalla residenza in una determinata area geografica. Pur esprimendo parere favorevole, l’Autorità ha suggerito di introdurre una serie di garanzie per il rispetto dei diritti degli interessati. In particolare, l’Autorità ha indicato di pubblicare l’informativa sul sito istituzionale, promuovere un contraddittorio con il cittadino e consentirgli di replicare alla valutazione fatta dall’algoritmo.

Nel settore sanitario, l’Autorità si è espressa sull’uso di trattamenti automatizzati a fini di *governance* sanitaria in un parere al Consiglio di Stato nel 2020 (17). Sulla base di un piano proposto dal Ministero della Salute, il Dicastero intendeva procedere alla distribuzione delle risorse economiche dello Stato tra le Regioni, passando da un modello basato sull’età della popolazione ad uno fondato sull’effettivo fabbisogno economico del territorio. Per raggiungere questo obiettivo, il Ministero aveva proposto di rendere interoperabili molteplici banche dati per tracciare il “profilo sanitario individuale” di ogni singolo utente del Ssn, da collegare a quello reddituale (il c.d. “*status sociale*”). La transizione al nuovo modello di *governance* sanitaria presupponeva dunque la profilazione dello stato di salute di ogni singolo assistito dal Ssn, per suddividere la popolazione in gruppi omogenei per condizione patologica e reddituale. Nel parere il Garante ha rilevato l’assenza di una base normativa, necessaria per raggiungere gli obiettivi ministeriali, richiamando l’attenzione sulla necessità che i trattamenti di dati personali connessi a tale nuovo sistema di ripartizione siano effettuati nel pieno rispetto della disciplina sulla protezione dei dati personali. Sulla scorta di alcuni precedenti e della nota sentenza del Consiglio di Stato del 2020 (18), il Garante ha applicato tre fondamentali principi: 1) la conoscibilità dei processi decisionali automatizzati; 2) la non esclusività della decisione algoritmica; e 3) la non discriminazione algoritmica.

In ambito giuslavoristico è meritevole di nota il provvedimento adottato nei confronti di una nota società di *food delivery*, Deliveroo Italy S.r.l. (19). All’esito di una complessa istruttoria avviata d’ufficio dall’Autorità, emergeva che i circa 8.000 fattorini impiegati presso Deliveroo subivano un’invasiva profilazione finalizzata a valutarne l’affidabilità e la disponibilità e assegnare loro gli ordini all’interno delle fasce orarie prenotate. Con il sistema algoritmico proprietario “Frank” veniva valutato l’operato del fattorino, senza che fosse realmente consapevole dei meccanismi di *feedback* e di *scoring* in atto. Il Garante comminava alla società un’ingente sanzione amministrativa pecuniaria, ingiungendo alla società di garantire ai *rider* di poter conoscere la logica algoritmica, esprimere la propria opinione, ottenere l’intervento di un addetto umano e contestare la decisione adottata. Veniva anche ordinata l’adozione di misure e procedure matematiche o statistiche per monitorare la correttezza e accuratezza algoritmica, per minimizzare il rischio di errori ed evitare usi impropri e discriminatori dei meccanismi reputazionali basati su *feedback*.

In un provvedimento articolato in materia di tutela dell’ordine pubblico, il Garante si pronunciava sul sistema mobile Sari Real Time (20) su richiesta del Ministero dell’Interno. Grazie all’installazione di videocamere in diverse aree geografiche, Sari Real Time avrebbe permesso l’analisi, in tempo reale, dei volti dei soggetti ripresi, confrontandoli con una banca dati contenente fino a 10.000 volti (“*watch-list*”). Se l’algoritmo di riconoscimento facciale avesse riscontrato una corrispondenza tra un volto ripreso dalla videocamera e un altro contenuto nella *watch-list* il sistema avrebbe generato un *alert* per richiamare l’attenzione degli operatori delle Forze dell’Ordine. Veniva riscontrata l’assenza di una idonea base giuridica a copertura di attività di trattamento automatizzato di dati biometrici per le dichiarate finalità di sicurezza. Sottolineata l’estrema delicatezza della questione, l’Autorità evidenziava il rischio che Sari Real Time avrebbe aperto a trattamenti automatizzati su larga scala idonei a incidere anche su persone che, semplicemente, si riunivano per manifestare il proprio convincimento politico o sindacale. La finalità originaria sottesa all’adozione di Sari Enterprise Real Time si sarebbe dunque evoluta, passando da una sorveglianza “individualizzata” a una sorveglianza “di massa” poiché l’acquisizione di nuovi volti avrebbe generato nuovi modelli di immagini da confrontare con quelli dei soggetti all’interno della *watchlist*. Pertanto, veniva reso parere negativo. Il sistema Sari Real Time non veniva giudicato conforme alla disciplina vigente, data l’assenza di adeguate e specifiche disposizioni normative legittimanti i trattamenti in questione. In un altro significativo provvedimento dello stesso anno, l’Autorità si pronunciava sul noto caso “Clearview AI” (21), cioè un motore di ricerca per il riconoscimento facciale offerto dall’omonima società statunitense. Con tecniche di *web scraping*, il titolare aveva raccolto immagini da *social network*, *blog* e *siti web* per poi estrarne le caratteristiche biometriche e conservarle nel proprio *database*. La banca dati così costituita conteneva oltre 10 miliardi di immagini di persone, arricchita da dati personali come la geolocalizzazione, la data di nascita ed il genere dell’individuo.

(16) Provv. n. 515 del 21 novembre 2013 (doc. web n. 2765110).

(17) Provv. n. 43 del 5 marzo 2020 (doc. web n. 9304455).

(18) Cons. Stato, Sez. VI, sent. n. 881 del 4 febbraio 2020.

(19) Provv. n. 285 del 22 luglio 2021 (doc. web n. 9685994).

(20) Provv. n. 127 del 25 marzo 2021 (doc. web n. 9575877).

(21) Provv. n. 50 del 10 febbraio 2022 (doc. web n. 9751362).

Si poteva quindi determinare a chi appartenesse quello specifico volto, dove era stata scattata la foto, quale era il contesto, e persino quanti anni aveva colui o colei che vi era ritratto, consentendo a *Clearview* la creazione di profili dettagliati di cittadini a loro insaputa. Eppure, è necessaria l'espressione del consenso per il trattamento di immagini biometriche, previa visione di idonea informativa. Pertanto, il Garante comminava una sanzione pecuniaria di 20 milioni di euro, vietava la prosecuzione del trattamento e l'ulteriore raccolta di dati personali con tecniche di *web scraping* di dati biometrici, oltre a ordinare la cancellazione dei dati comuni e biometrici raccolti e trattati.

Con il noto provvedimento di limitazione provvisoria nei confronti di ChatGPT (22), il Garante metteva in evidenza diverse criticità legate all'uso di sistemi di c.d. IA generativa. Innanzitutto, veniva riscontrata l'assenza di un'idonea base giuridica in relazione alla raccolta dei dati personali e al loro trattamento per scopo di addestramento degli algoritmi sottesi al funzionamento di ChatGPT. Inoltre, il trattamento di dati personali degli interessati risultava inesatto, perché gli *output* generati dal modello non sempre risultavano corrispondenti al dato reale. Non era stato nemmeno adottato alcun meccanismo di *age verification* in relazione al servizio ChatGPT che, secondo i termini allora pubblicati da OpenAI L.L.C., era riservato a soggetti di almeno 13 anni. Pertanto, i minori di età inferiore a 13 anni rischiavano di restare esposti a contenuti inidonei rispetto al loro grado di sviluppo psicofisico. Per queste ragioni veniva ravvisata la necessità di disporre in via d'urgenza la misura della limitazione provvisoria del trattamento, estesa al territorio italiano.

3. La tutela della dignità personale all'interno dell'AI Act, tra prassi di IA vietate e "ad alto rischio".

Con l'AI Act, l'Unione europea ha scelto di orientare lo sviluppo dell'IA a promozione dei diritti e della dignità umana tramite una regolamentazione attenta. La diffusione capillare e trasversale dei sistemi di IA ha richiesto al legislatore europeo un atto di coraggio: tracciare una linea di confine tra prassi ammissibili e usi vietati o ad alto rischio dell'IA. (23).

Nell'AI Act la promozione della dignità personale viene declinata in negativo, tramite l'imposizione di divieti di determinati usi dell'IA, o nella previsione di specifici obblighi positivi qualora gli utilizzi dell'IA ricadano nell'"elevato rischio" definito dal Regolamento.

L'art. 5 dell'AI Act vieta diverse pratiche di utilizzo dell'IA. La *ratio* delle proibizioni imposte si ritrova nel valore fondante dell'Unione europea di dignità personale, nonché di tutela dei diritti fondamentali sanciti dalla CDFUE, compresi il diritto alla non discriminazione, alla protezione dei dati e alla vita privata e i diritti dei minori. Il regolatore è consapevole che l'IA presenta, oltre a un vasto potenziale benefico, il rischio di essere utilizzata impropriamente e di fornire strumenti nuovi e potenti per pratiche di manipolazione, sfruttamento e controllo sociale (24). Si pensi al divieto sancito dall'art. 5, par. 1, lett. a) (25) di usi dell'IA subliminali, manipolativi o ingannatori, con lo scopo di orientare il comportamento delle persone, inducendole a compiere scelte che altrimenti non avrebbero mai preso. O ancora, alla proibizione prevista dall'art. 5, par. 1, lett. c), (26) di usare l'IA per finalità di c.d. *social scoring*, ovvero per valutare o classificare le persone sulla base del loro comportamento, caratteristiche, personalità per determinare un punteggio e danneggiarle o sfavorirle nella società. Se l'esperienza cinese in questo senso è paradigmatica (27), non si possono sottovalutare sviluppi o utilizzi dell'IA per finalità di c.d. *social scoring* anche all'interno del territorio europeo e nazionale. Si pensi al caso di diversi enti locali che avevano avviato progetti sperimentali di *social scoring* (28), che hanno reso necessario l'intervento dell'Autorità Garante determinando l'avvio di tre istruttorie. Se il cittadino bisognoso avesse mantenuto comportamenti non ritenuti "in linea" dall'amministrazione comunale, avrebbe perso punteggio e rischiato di perdere il diritto all'alloggio. Paradigmatico, pur con le sue numerose e ampie deroghe, il divieto disposto dall'art. 5, par. 1, lett. h) (29)

(22) Provv. 30 marzo 2023, n. 112 (doc. web n. 9870832).

(23) P. Stanzione, *Regolare il futuro: La protezione dei dati per un'innovazione antropocentrica*. Relazione del Presidente dell'Autorità Garante per la Protezione dei Dati Personali, Roma, 3 luglio 2024: "La continua espansione ed evoluzione dell'intelligenza artificiale impone dunque di tracciare (e questo è il massimo compito della politica) un limite di sostenibilità, delle colonne d'Ercole da non varcare perché il progresso non divenga, paradossalmente, socialmente regressivo".

(24) Cfr. il cons. 28 dell'AI Act.

(25) Cfr. art. 5, par. 1, lett. a), AI Act.

(26) Cfr. art. 5, par. 1, lett. c), AI Act.

(27) Per un approfondimento sulla diffusione dei sistemi di IA di c.d. *social scoring* in Cina, cfr. I.A. Filipova, *Legal Regulation of Artificial Intelligence: Experience of China*, in *Journal of Digital Technologies and Law*, vol. 2, n. 1, 46-73, 2024; C. Li, H. Wang, S. Jiang, B. Gu, *The Effect of AI-Enabled Credit Scoring on Financial Inclusion: Evidence from One Million Underserved Population*, in *SSRN Electronic Journal*, 2024; N. Raghunath, *Theorising quantified credibility in the age of big data: a case of China's Social Credit System*, in *The Journal of Chinese Sociology*, vol. 11, n. 1, 2024; W. Z. Hu, *Understanding the Power of China's National Social Credit System: A Structural/Mechanism Explanation*, in *Philosophy of the Social Sciences*, vol. 54, n. 3, 203-225, 2024; View; Y. Chen, A.S. Cheung, *The Transparent Self Under Big Data Profiling: Privacy and Chinese Legislation on the Social Credit System*, Working Paper, University of Hong Kong, 2017.

(28) Autorità Garante per la Protezione dei Dati Personali, "Cittadinanza a punti": Garante privacy ha avviato tre istruttorie. Preoccupanti i meccanismi di *scoring* che premiano i cittadini "virtuosi", 8 giugno 2022, doc. web n. 9778361.

(29) Cfr. art. 5, par. 1, lett. h), AI Act.

dell'uso dei sistemi di IA di identificazione biometrica remota “in tempo reale” in spazi accessibili al pubblico a fini di tutela dell'ordine pubblico e repressione dei reati. In tal senso, è stata formativa l'esperienza italiana del progetto “SARI” del Ministero dell'interno. Quel progetto di sorveglianza biometrica, poiché privo di un'idonea base giuridica a legittimarlo, avrebbe conferito ampi margini di discrezionalità agli operatori delle forze dell'ordine, aprendo a inedite forme di sorveglianza di massa.

Seguendo l'approccio c.d. *risk-based*, l'art. 6 dell'*AI Act* classifica diversi usi dell'IA come rientranti nella categoria “ad alto rischio” perché ritenuti capaci di incidere significativamente sulla dignità personale e su diversi diritti fondamentali dell'individuo (30). Tra questi, il legislatore europeo annovera il diritto alla protezione dei dati personali, oltre al diritto alla non-discriminazione, alla salute, all'istruzione e ai diritti dei lavoratori, dei minori, delle persone con disabilità e di altri soggetti vulnerabili (31). Tali usi ad “alto rischio” sono individuati in relazione al rischio concreto di pregiudizio per la salute e la sicurezza dei consumatori, alla possibilità di profilare gli individui e rispetto a specifici settori di incidenza elencati nell'all. 3 all'*AI Act*, come “istruzione e formazione professionale”, “occupazione, gestione dei lavoratori e accesso al lavoro autonomo”, “gestione delle migrazioni e controllo delle frontiere”, “accesso ai servizi pubblici e privati essenziali”.

La qualificazione di un sistema di IA come “ad alto rischio” impone concretamente a fornitori (32), rappresentanti autorizzati dei fornitori (33), importatori (34), distributori (35) e *deployer* (36) l'obbligo di rispettare una serie di condizioni per l'immissione sul mercato o la messa in servizio di tali sistemi, oltre all'osservanza di requisiti specifici (37). Tra questi ultimi, si pensi all'obbligo dei *deployer* di svolgere una valutazione di impatto sui diritti fondamentali degli individui (38). O all'obbligo gravante sui fornitori di addestrare i sistemi di IA con dati rappresentativi della popolazione, completi ed accurati, con l'obiettivo di evitare distorsioni e *bias*, tramite un'adeguata *governance* dei dati (39). Ancora, l'*AI Act* sancisce il principio di supervisione umana (40), che si traduce concretamente nell'obbligo in capo ai fornitori di progettare i sistemi di IA in modo tale da assicurare sempre all'interessato il controllo di un addetto umano competente e qualificato durante il suo impiego e, se del caso, l'arresto del funzionamento in caso di anomalie e problemi. Si pensi anche al diritto di ciascuna persona interessata dalla decisione adottata dall'utilizzatore di un sistema di IA, che produca effetti giuridici o in modo analogo incida significativamente sulla sua persona, realizzando un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi diritti fondamentali, di ottenere spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata (41). In concreto, questo diritto si sostanzia nella possibilità per chi è interessato dalla decisione assunta dal c.d. *deployer* di ottenere spiegazioni “chiare e significative” sul ruolo assunto dall'IA nel procedimento decisionale automatico e sui principali elementi costitutivi.

(30) Cfr. cons. 7, *AI Act*.

(31) Cfr. cons. 1, 3, 5, 7, 8, 48, *AI Act*.

(32) Oltre ai requisiti specifici dei sistemi di IA ad alto rischio, sui fornitori gravano differenti obblighi, come fornire i propri dati di contatto, istituire un sistema di gestione della qualità, conservare la documentazione inerente al sistema di IA e i relativi log, mantenere aggiornata una dichiarazione di conformità, apporre la marcatura CE, rispettare i vincoli di registrazione nella banca dati europea, conformarsi agli standard di accessibilità, monitorare il funzionamento dei sistemi dopo la loro immissione sul mercato e segnalare qualsiasi incidente grave alle Autorità competenti. Cfr. artt. 16, 17, 18, 19, 20, 43, 48, 49, 72, *AI Act*.

(33) Ai sensi dell'art. 11 dell'*AI Act*, i rappresentanti autorizzati dei fornitori devono aver ricevuto una nomina ufficiale dal fornitore, verificare la presenza della dichiarazione di conformità Ue, della documentazione tecnica inerente al sistema di IA e lo svolgimento della procedura di valutazione della conformità, conservare tale documentazione per un periodo di dieci anni dall'immissione e messa in servizio del sistema di IA, nonché rispettare gli obblighi di registrazione alla Banca dati Ue.

(34) Gli importatori dei sistemi di IA sono tenuti a svolgere controlli di conformità prima della loro immissione sul mercato ed eventualmente sospenderla in presenza di discrasie documentali, garantire che le condizioni di stoccaggio o trasporto non ne pregiudichino la conformità e conservare la relativa documentazione per almeno dieci anni dall'immissione sul mercato o messa in servizio. Cfr. l'art. 23, *AI Act*.

(35) I distributori dei sistemi di IA devono verificare in via preliminare il rispetto dei requisiti di conformità all'*AI Act* da parte del fornitore e dell'importatore ed eventualmente sospenderne l'immissione sul mercato, garantire idonee condizioni di stoccaggio o trasporto e segnalare al fornitore o all'importatore qualsiasi rischio sopraggiunto. Cfr. l'art. 24 dell'*AI Act*.

(36) I *deployer* di un sistema di IA ad alto rischio sono tenuti a rispettarne le istruzioni d'uso, garantire la sorveglianza umana, assicurare la qualità e rappresentatività dei dati di input, nonché segnalare rischi notevoli o gravi incidenti occorsi agli altri operatori coinvolti e all'Autorità di vigilanza. Cfr. l'art. 26 dell'*AI Act*.

(37) I fornitori dei sistemi di IA ad alto rischio sono tenuti a diversi adempimenti specifici. Cfr. artt. 8, 9, 10, 11, 12, 13, 14 e 15 dell'*AI Act*.

(38) Cfr. art. 27 e cons. 28, *AI Act*.

(39) Cfr. art. 10, *AI Act*.

(40) Cfr. art. 14, *AI Act*.

(41) Cfr. art. 86, par. 1, *AI Act*.

4. La sfida della governance in materia di IA. Conclusioni e prospettive future.

Sia nella sistematica del Gdpr che dell'*AI Act*, la protezione della dignità umana è un *fil rouge*. L'effettività della sua tutela dipende dall'azione di controllo e vigilanza delle Autorità competenti individuate dai Regolamenti.

Con il Gdpr il legislatore europeo ha affidato la competenza ad Autorità amministrative indipendenti, entità caratterizzate da indipendenza, imparzialità e terzietà nell'esercizio dei relativi compiti ed attività. Tale scelta regolatoria è andata nella direzione di istituire dei reali presidi di tutela del diritto fondamentale alla protezione dei dati personali. Invece nell'*AI Act* è stata imboccata una strada differente, problematica nell'ottica di garantire l'effettività della tutela della dignità umana e dei diritti fondamentali dell'individuo. Nonostante l'incidenza della materia dell'IA sull'esercizio di diritti e libertà fondamentali, il novero dei soggetti destinatari di competenze di supervisione non risulta circoscritto ad Autorità amministrative indipendenti, pur esigendo che ne sia salvaguardata l'obiettività e l'imparzialità nell'esercizio dei relativi compiti ed attività (42).

Nel testo dell'*AI Act*, il ruolo di vigilanza delle Autorità garanti per la protezione dei dati personali appare critico, ma non chiaramente definito a livello operativo in alcuni settori strategici. Tra questi rientra una competenza cruciale come la supervisione sui sistemi di IA ad alto rischio usati a fini di attività di contrasto, gestione delle frontiere, giustizia e democrazia (43). Si tratta di un profilo non irrilevante, nella misura in cui determina un sostanziale affievolimento delle garanzie dei diritti fondamentali e, in particolare, del diritto alla protezione dei dati personali, la cui tutela effettiva si fonda sulla necessaria vigilanza delle Autorità di controllo ai sensi degli artt. 8 CDFue e 16 Tfue. Oltre a ciò, il ruolo delle Autorità è limitato a compiti minori come il coinvolgimento all'interno delle *sandbox* regolamentari (44), la ricezione di notifiche sull'uso dei sistemi di identificazione biometrica da remoto "*real time*" (45) e la redazione di una relazione annuale su tali usi da trasmettere alla Commissione europea (46).

Seppur l'*AI Act* sancisca espressamente una clausola di salvaguardia a favore della disciplina in materia di protezione dati personali in relazione ai dati personali trattati in relazione ai diritti e agli obblighi stabiliti dall'*AI Act* (47), è evidente una mancanza di coordinamento tra questi plessi normativi. Non si tratta di speculazioni teoriche, ma preoccupazioni reali in quanto i problemi di coordinamento normativo non si risolvono con la previsione di formule di salvaguardia. Sarebbe stato auspicabile prevedere nel testo dell'*AI Act* degli efficaci meccanismi di coordinamento operativo tra le diverse Autorità coinvolte nel ciclo di vita dei sistemi e dei modelli di IA, nel rispetto del principio di leale cooperazione. In loro assenza, emergono importanti lacune e discrasie nella complessiva struttura di *governance* dell'IA, che si traducono in una menomazione delle competenze affidate alle Autorità garanti di protezione dei dati personali (48).

Di fronte all'impatto dirompente dell'IA sulla tutela della dignità umana e dei diritti fondamentali della persona, pare indifferibile la designazione delle Autorità di protezione dei dati personali come Autorità competenti per la vigilanza sui sistemi di IA ad alto rischio usati per le attività di contrasto, gestione delle frontiere, giustizia e democrazia nel rispetto dell'art. 74, par. 8 dell'*AI Act* (49). In prospettiva, pare inoltre auspicabile affiancare alle Autorità di notifica e vigilanza del mercato designate ai sensi dell'art. 70 dell'*AI Act* l'Autorità garante in materia di protezione dei dati personali designata come "*Autorità garante dei diritti fondamentali*" ai sensi dell'art. 77 dell'*AI Act* (50), affinché vigili sull'incidenza dell'IA sui diritti umani esercitando i poteri riservatele da tale disposizione. Si tratta di poteri di richiesta informativa ed accesso documentale alla documentazione formata in relazione all'uso di sistemi di IA ad alto rischio necessari da esercitare per adempiere efficacemente ai propri compiti istituzionali (51). A titolo esemplificativo, l'Autorità garante in materia di protezione dati personali,

(42) Cfr. art. 70, par. 1, *AI Act*.

(43) Cfr. art. 74, par. 8, *AI Act*.

(44) Cfr. art. 57, par. 10, *AI Act*.

(45) Cfr. art. 5, par. 4 e 5, *AI Act*.

(46) Cfr. art. 5, par. 6, *AI Act*.

(47) Cfr. art. 2, par. 7, *AI Act*.

(48) Come rilevato di recente in Comitato europeo di protezione dei dati personali, *Statement 3/2024 on data protection authorities' role in the Artificial Intelligence Act framework*, 16 luglio 2024, par. 11: "From a broader perspective, there is a need for sound cooperation between MSAs and the other entities which are tasked with the supervision of AI systems²⁰, including DPAs, and clear procedures have to be provided for in this regard on the basis of Article 74(10) AI Act²¹. Such procedures should be built and developed under the principle of sincere cooperation provided by the Article 4(3) of the Treaty of the European Union, as highlighted by the Court of Justice in the Bundeskartellamt case. In this way, inconsistencies between decisions taken by different oversight authorities and bodies can be prevented in the digital ecosystem, and synergies can be exploited in coherent, effective and complementary enforcement actions for the benefit of individuals and legal certainty".

(49) Cfr. Comitato europeo di protezione dei dati personali, *Statement 3/2024 on data protection authorities' role in the Artificial Intelligence Act framework*, 16 luglio 2024.

(50) Sia consentito il rinvio a G. Cerrina Feroni, Commissione straordinaria per la tutela e la promozione dei diritti umani: Audizione della prof.ssa Ginevra Cerrina Feroni, Senato della Repubblica, 9 maggio 2024.

(51) Cfr. art. 77, par. 1 dell'*AI Act*.

se fosse designata ai sensi dell'art. 77 dell'*AI Act*, in sede di istruttoria per accertare una possibile violazione del Gdpr potrebbe accedere alla documentazione formata dal titolare del trattamento per rispettare gli obblighi dell'*AI Act*, come quella relativa alle misure di gestione dei rischi, ai *dataset* di addestramento, convalida e prova dei sistemi di IA, alla valutazione di impatto sui diritti fondamentali e così via. Diversamente, il diritto fondamentale alla protezione dei dati personali rischierebbe di diventare un guscio vuoto, privo di effettività, a detrimento dell'intangibilità della dignità personale.

* * *

PER UNA GENEALOGIA CRITICA DELL'INTELLIGENZA ARTIFICIALE

di Teresa Numerico (*)

Abstract: L'articolo ha l'obiettivo di effettuare una genealogia dell'attuale contesto di successi dell'Intelligenza artificiale. Il testo offre una descrizione di tutte le alternative concettuali che ne hanno animato la storia lunga settant'anni. Dalla versione dell'IA di impostazione logicista, alla cibernetica, dal connessionismo, all'integrazione con la macchina, fino ad arrivare ai risultati più promettenti che si sono succeduti negli ultimi dieci anni che derivano da alcuni metodi che si basano sull'addestramento offerto da ingenti corpora di dati, con la finalità di estrarre regolarità e schemi noti come pattern da proiettare sulle serie future dei medesimi dati per prevederne l'orientamento e favorire l'intervento. Queste strategie note come machine learning e deep learning hanno ottenuto notevoli risultati, ma le loro applicazioni necessitano di una chiara funzione obiettivo. Senza di essa non abbiamo modo di valutarne gli effetti a causa dell'impossibilità di controllare i processi interpretativi e l'organizzazione del calcolo delle probabilità, dal momento che i dati sono troppi per controllarli a mano. Inoltre, questi metodi basati sull'apprendimento e l'addestramento attraverso i dati conducono a dei risultati che tendono alla standardizzazione e alla normalizzazione con la tendenza ad amplificare gli stereotipi già presenti nei corpora. Tale effetto è particolarmente eclatante nell'ambito dell'IA generativa. La produzione artificiale di contenuti non può nemmeno garantire la loro veridicità perché ciò implicherebbe una curatela dei dati di addestramento, impossibile considerata la loro enorme quantità. Chi usa questi strumenti è costretto a controllare i risultati, verificando le risposte per evitare le allucinazioni, o meglio gli errori che sistematicamente possono verificarsi, in modo imprevisto, ma costante. Questo dovrebbe farci riflettere sulle scelte politiche relative a quali compiti possano essere svolti da questi sistemi in modo autonomo e quali richiedano invece l'integrazione stretta con le capacità umane e quali debbano essere assunti esclusivamente da esseri umani, anche in relazione alle problematiche di responsabilità delle decisioni esercitate.

The article aims to perform a genealogy of the current context of success surrounding Artificial Intelligence. It offers a description of all the conceptual alternatives that animated its history over the past seventy years. From the logicist version of AI to cybernetics, from connectionism to integration with machines, it explores the most promising results that emerged over the last decade, derived from methods based on training with large data corpora, aimed at extracting patterns to project onto future series of the same data for prediction and intervention purposes. These strategies, known as machine learning and deep learning, have achieved significant results; however, their applications require a clear objective function. Without it, we cannot evaluate their effects due to the impossibility of controlling interpretative processes and the organization of probability calculations, given that the data volume is too large for manual control. Moreover, these data-driven learning and training methods lead to outcomes that tend toward standardization and normalization, amplifying existing stereotypes within the corpora. This effect is particularly pronounced in the realm of generative AI. The artificial production of content cannot guarantee its truthfulness, as this would imply curating the training data an impossible task, given their vast quantity. Users of these tools are compelled to verify results, checking responses to avoid hallucinations or systematic errors that can occur unexpectedly but consistently. This should prompt us to reflect on political choices regarding which tasks can be autonomously performed by these systems, which require close integration with human capabilities, and which should be exclusively undertaken by humans, especially in cases in which there are decision-making processes, that require a clear accountability structure.

Sommario: 1. Introduzione. – 2. Quando è nata l'intelligenza artificiale? – 3. L'intelligenza artificiale ha molte anime. – 4. La cibernetica e la comunicazione con la macchina. – 5. Tra reasoning e learning: diverse strategie al confronto. – 6. Machine learning. – 7. La centralità dei dati. – 8. Criteri di validazione della conoscenza. – 9. Intelligenza artificiale generativa: la creatività dei modelli. – 10. La politica dei sistemi intelligenti.

(*) T. Numerico è professoressa associata di Logica e Filosofia della Scienza all'Università di Roma Tre.

Alcune parti di questo articolo sono una rielaborazione del primo capitolo del volume *Big Data e algoritmi*, Carocci, Roma, 2021.

1. Introduzione

Ad attribuire l'intelligenza alla macchina tra i pionieri dell'informatica aveva pensato Alan Turing (1). Il suo suggerimento, perché i dispositivi fossero accreditati come intelligenti, era che fingessero di maneggiare il linguaggio. Era l'obiettivo esplicitamente assunto dal suo famoso Test, o gioco dell'imitazione. Dalla sua attenzione per il linguaggio e per la sua difficile governabilità da parte della macchina si passò a una visione riduzionista dell'IA secondo cui le macchine sarebbero in grado di sostituire la mente attraverso un sistema fisico-simbolico di impostazione logicista. Contemporaneamente a questo paradigma della simulazione della mente, si sviluppa la cibernetica, un sapere transdisciplinare, il cui obiettivo è abbattere la differenza tra organico e inorganico attraverso l'analisi dell'attività degli agenti, considerati nelle loro prerogative di comunicazione e controllo. Il modello di integrazione tra vivente e macchina conduce a immaginare l'intelligenza meccanica come l'esito di un processo di progressivo ampliamento delle capacità umane, attraverso strategie di comunicazione e controllo.

La prospettiva dell'interazione uomo-macchina si è evoluta negli ultimi anni in una tendenza alla registrazione in forma di dati dell'azione umana, possibile grazie alla diffusione della rete Internet che ha consentito una registrazione di tutte le azioni compiute nell'ambito della comunicazione digitale. L'intelligenza artificiale si è trasformata, così, nella costruzione di algoritmi di apprendimento in grado di interpretare i dati ai fini di prevedere, anticipare e orientare il comportamento umano, dopo averlo discretizzato in un flusso di dati misurabili. Non si tratta più di una simulazione dell'attività umana, ma di una sua riproduzione, dopo averla reificata nella forma delle sue tracce online. La replica digitale serve per anticipare abitudini e preferenze delle persone, e per valutare e orientare le loro attività. La riproduzione che gli algoritmi di *machine learning* e *deep learning* prospettano riguarda sia l'essere umano come oggetto di analisi, sia l'essere umano in quanto decisore e interprete delle informazioni disponibili.

Il computer, infatti, a queste condizioni, svolge il ruolo per il quale è più adatto: quello di una macchina per eseguire calcoli. Dopo aver cercato, quindi, invano di esorcizzare il corpo chiudendolo tra parentesi, nell'ipotesi di riuscire a costruire sistemi simbolici del tutto privi di riferimento fisico, attraverso la mediazione della cibernetica, l'intelligenza artificiale si è riorientata verso l'interazione comunicativa tra esseri umani e macchine. Tale interazione in un primo tempo si è manifestata come integrazione in cui la macchina era sotto il controllo dell'operatore umano e poi, con l'avvento dei dati prodotti dalle azioni umane registrate, ha sussunto sotto di sé la capacità di prendere decisioni e quella di generare autonomamente contenuti, sia pure dipendendo, per ora, dall'operatore umano per la costruzione degli algoritmi per interpretare i dati. Tale evoluzione propone una serie di questioni di natura epistemologica, etica, sociale e politica che vale la pena analizzare.

La prima questione riguarda la presa di decisione. Non è facile, infatti, sottoporre al controllo di operatori umani le decisioni algoritmiche. Come suggeriva preoccupato Wiener, gli esseri umani rischiano di essere lasciati fuori dal *feedback* perché troppo lenti (2).

La seconda questione è che la macchina, come segnalava Turing (3), deve solo fingere di essere intelligente ed essere capace di ingannare giudici non esperti di tecnologia. Ma se non si può controllare quello che il dispositivo fa – e il meccanismo è programmato per ingannare gli umani – come possiamo fidarci delle sue decisioni?

2. Quando è nata l'intelligenza artificiale?

Il termine Intelligenza Artificiale è ormai diventato imprescindibile. Qualsiasi attività sembra poter essere aumentata, migliorata o sostituita utilizzando i suoi sistemi. Quando l'espressione fu coniata – come titolo del progetto del 1956 per un *Summer workshop* del Dartmouth College – il concetto non era ancora famoso (4). Il gruppo che firmò la proposta fu coordinato da John McCarthy, che nel 1955 richiese il finanziamento di un seminario ristretto a una decina di partecipanti alla Rockefeller Foundation. La scelta del termine avvenne proprio perché non si riferiva precisamente a nessuna delle ricerche in voga in quegli anni tra la cibernetica e la teoria degli automi. Diversamente dall'espressione meno impegnativa usata precedentemente da Alan Turing nei suoi articoli

(1) Cfr. Turing A. (1950), *Computing machinery and intelligence*, "Mind", vol. 59, 433-460; ristampato in Copeland (2004), *The essential Turing*, Clarendon Press, Oxford, 441-464 (trad.it. *Macchine calcolatrici e intelligenza*, in Lolli G. (a cura di) *Intelligenza Meccanica*, Bollati Boringhieri, Torino, 1994, 121-57).

(2) Cfr. Wiener N. (1950/1954), *The Human Use of Human Beings*, Houghton Mifflin, Boston (trad. it. *introduzione alla cibernetica: l'uso umano degli esseri umani*, Bollati Boringhieri, Torino, 1966) e Wiener N. (1960), *Some Moral and Technical Consequences of Automation*, in "Science", New Series, vol.131(3410) (May 6, 1960), pp. 1355-8.

(3) Cfr. Turing A.M. (1948/2004), *Intelligent machinery*, in Copeland J. (Ed.) (2004), *The essential Turing*, Clarendon Press, Oxford, pp.410-432 (trad.it. *Macchine intelligenti*, in Lolli G. (a cura di) *Intelligenza Meccanica*, Bollati Boringhieri, Torino, 1994, pp.88-120) e Turing A.M. (1952/2004), *Can Automatic calculating machines be said to think?*, discussione registrata alla BBC il 10 gennaio 1952 tra Turing, Max Newman, Richard Braithwaite e Geoffrey Jefferson, pubblicata in Copeland J. (Ed.) (2004), *The essential Turing*, Clarendon Press, Oxford, pp.494-506.

(4) McCarthy J., Minsky M. Rochester N. Shannon C. (1955), *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, in "AI Magazine", vol.27 Number 4 (2006), 12-4; consultato il 29 settembre 2024) (trad. it. *Proposta di un Progetto di ricerca estivo sull'intelligenza artificiale presso il Dartmouth College*, in *Sistemi intelligenti*, XVIII, dic. 2006, 413-428).

sul tema (5) e cioè macchine intelligenti (*intelligent machinery*), *artificial intelligence* si riferisce alla possibilità che queste macchine di nuova creazione possano simulare l'intelligenza umana, non solo eseguire procedure considerate intelligenti.

Il progetto di McCarthy era ambizioso, ma inclusivo teneva insieme diverse attività di ricerca, come la capacità di simulazione di compiti intelligenti, le regole per la programmazione di un linguaggio in grado di eseguire manipolazioni simboliche, la creazione delle cosiddette reti neurali, che propendono per modellare il funzionamento delle attività cerebrali, piuttosto che dell'intelligenza *tout court*, i metodi di autoapprendimento delle macchine, la capacità di astrazione e creatività, oltre che i metodi per misurare la complessità di un calcolo.

Nonostante la fase embrionale nella quale le ricerche si trovavano, nel progetto iniziale emergevano già alcune delle grandi dicotomie che hanno caratterizzato la storia burrascosa di questa disciplina che – pur essendo giovane (non sono ancora 70 anni da quando il termine è stato coniato) – ha ricevuto moltissime attenzioni scientifiche e mediatiche, attraversando varie 'estati' e vari 'inverni'.

3. L'intelligenza artificiale ha molte anime

La prima contrapposizione riguarda l'oggetto della simulazione, cioè se le macchine debbano simulare il funzionamento delle capacità cognitive di alto livello, come la capacità di manipolare simboli e ottenere risultati nell'ambito della matematica, del *problem solving* o di altre discipline, formali e astratte, o se invece debba simulare o modellare il funzionamento cerebrale, molto più complesso, ma indipendente da meccanismi di manipolazione simbolica comprensibili.

La seconda opposizione metodica e contenutistica riguarda il ragionamento opposto all'apprendimento (*reasoning vs learning*). Se si debbano concentrare le forze a simulare il ragionamento, ovvero la capacità di trarre conclusioni affidabili da premesse prestabilite in base a regole di inferenza definite prima di cominciare a usarle, o se sia più produttivo e vantaggioso partire dai dati disponibili e cercare metodi per estrarre regolarità e modelli dinamici di funzionamento del fenomeno in esame. In questo secondo caso la macchina lavorerebbe su strategie di apprendimento, costruite dal basso verso l'alto, cioè a partire dai dati oggetto di investigazione, senza avere in anticipo una rappresentazione simbolica del problema da risolvere. Ci sono vantaggi e svantaggi in ognuna delle soluzioni. Il ragionamento permette di spiegare i risultati ottenuti dalla macchina, perché le sue regole sono esplicite, ma non riesce ad andare oltre il programma, e rischia facilmente di incorrere in una serie di potenziali esplosioni computazionali che rendono impossibile ottenere un risultato per problemi del mondo reale. L'apprendimento invece può fornire sempre risultati utili, ma non è sempre affidabile nel dare conto delle ragioni di una certa conclusione. Possiamo, quindi, ottenere risultati inattesi e sorprendenti, senza essere sicuri della loro correttezza e spiegabilità. Inoltre, i dati usati per addestrare i programmi di apprendimento sono parte integrante della loro capacità di trarre conclusioni, ma conservano una certa contingenza o, peggio, una traccia dei meccanismi rappresentativi iniqui che caratterizzano le nostre società, specialmente se si riferiscono alle soggettività umane e alle loro fragilità.

La terza dicotomia riguarda gli obiettivi da attribuire alle prestazioni delle macchine eventualmente intelligenti, l'alternativa allude alla possibilità di usare le loro attività per acquisire nuove conoscenze o per esercitare abilità creative, sorprendenti ma non sempre affidabili rispetto alla descrizione del mondo a cui si riferiscono.

La quarta alternativa riguarda la scelta di usare le macchine per sostituire gli umani, o per amplificarne le capacità, integrandole con quelle macchiniche. Se vogliamo sostituire l'attività umana ci immaginiamo dispositivi, dotati di autonomia organizzativa, ai quali delegare tutto il processo cognitivo; se invece pensiamo a strumenti di amplificazione e integrazione, immaginiamo delle interfacce che ci permettano di comunicare agevolmente, ma sulle quali non si possa scaricare la responsabilità delle decisioni operative e che funzionino solo in relazione a richieste e input governati dagli esseri umani. La premessa e la rilevanza di questi dispositivi è connessa con le eventuali capacità comunicative e protesiche rispetto alle attività da gestire e il loro design presuppone il controllo umano dell'azione.

In questo articolo ci concentreremo soprattutto sulla dicotomia tra *reasoning* e *learning*, oltre che sulla distinzione tra integrazione e sostituzione che ha attraversato il campo delle macchine intelligenti o delle macchine per l'amplificazione dell'intelligenza, per poi descrivere alcune delle strategie di maggior successo dell'Intelligenza artificiale basata sul *Machine Learning*, che sono al centro del grande successo che ha caratterizzato l'ultimo decennio.

4. La cibernetica e la comunicazione con la macchina

Il progetto dell'intelligenza artificiale era affiancato fin dal principio da un altro modello di sviluppo delle macchine, quello della cibernetica, un campo di ricerche transdisciplinare, inventato da Norbert Wiener. La cibernetica considerava le macchine come dispositivi indistinguibili da altri agenti organici. Il titolo del testo del 1948 che diede vita a questa transdisciplina era: *Cybernetics or Control and Communication in the Animal and*

(5) Cfr. Turing, A. *Computing Machinery*, op.cit.

the Machine (6). Il suo obiettivo era analizzare i dispositivi sotto il profilo della comunicazione e del controllo. Vi si rappresentava il controllo come un tipo particolare di comunicazione in cui era necessario non solo trasmettere un messaggio informativo, ma anche essere sicuri che l'altro agente avesse ricevuto ed eseguito l'ordine. Secondo tale rappresentazione esseri viventi e dispositivi meccanici potevano essere analizzati attraverso lo stesso modello teorico di comportamento.

Il modello si basava sul *feedback* che permetteva all'agente che riceveva il messaggio di rispondere adattando il proprio comportamento al messaggio ricevuto dall'altro agente o dall'ambiente stesso. Vista da questa prospettiva, la capacità degli esseri viventi di adeguarsi all'ambiente per sopravvivere poteva essere valutata in modo simile a quella delle macchine di obbedire agli ordini impartiti o, meglio, di comportarsi in maniera teleologica, essendo in grado di portare a termine lo scopo che veniva loro prefissato.

Il principale modello di riferimento della cibernetica era la recente esperienza bellica nella quale radar, missili balistici e altri dispositivi meccanici funzionavano attraverso la fissazione di un obiettivo balistico, mostrando così di possedere una loro finalità nel raggiungere lo scopo che era loro attribuito, qualunque fosse: dal riconoscere un velivolo nemico sul radar, a inseguire un obiettivo in movimento per colpirlo. Si trattava dell'anticipazione di missili, droni e bombe intelligenti che hanno in quel frangente la loro origine.

Wiener came to see the predictor as a prototype not only of the mind of an inaccessible Axis opponent but of the Allied anti-aircraft gunner as well, and then even more widely to include the vast array of human proprioceptive and electro-physiological feedback systems. The model then expanded to become a new science known after the war as "cybernetics," a science that would embrace intentionality, learning, and much else within the human mind (7).

La tesi di Peter Galison – uno studioso di cibernetica membro della *Stanford School* (8) – è che la guerra avesse trasformato ogni nemico in un nemico meccanico:

On the mechanized battlefield, the enemy was neither invisible nor irrational; this was an enemy at home in the world of strategy, tactics, and maneuver, all the while thoroughly inaccessible to us, separated by a gulf of distance, speed, and metal. It was a vision in which the enemy pilot was so merged with machinery that (his) human-nonhuman status was blurred. In fighting this cybernetic enemy, Wiener and his team began to conceive of the Allied anti-aircraft operators as resembling the foe, and it was a short step from this elision of the human and the nonhuman in the ally to a blurring of the human-machine boundary in general (9).

Se il nemico non era un altro essere umano, ma un agente strettamente connesso con una macchina allora, in un certo senso, anche le forze aeree alleate erano simili al nemico. Sul campo di battaglia avveniva questa elisione, questa confusione del confine tra essere umano e macchina in generale.

Tale sconfinamento della macchina nell'ambito del vivente era stato già suggerito da un articolo scritto da Wiener con altri due suoi colleghi, Arturo Rosenblueth e Julian Bigelow, dal titolo *Behavior, purpose, teleology* (10), proprio durante la Seconda guerra mondiale.

A further comparison of living organisms and machines leads to the following inferences. The methods of study for the two groups are at present similar. Whether they should always be the same may depend on whether or not there are one or more qualitatively distinct, unique characteristics present in one group and absent in the other. Such qualitative differences have not appeared so far. The broad classes of behavior are the same in machines and in living organisms (Rosenblueth, Wiener, Bigelow 1943, 21).

Nell'articolo si mettevano le basi per affermare che ci fossero differenze di rilievo nei metodi di studio relativi a esseri viventi e alle macchine dal punto di vista della comunicazione e del controllo. Tale conclusione, quindi, prometteva un abbattimento sostanziale della barriera che separava organico e inorganico. La prospettiva di un possibile studio congiunto di esseri viventi e macchine rispetto alla comunicazione – di cui il controllo era solo un caso particolare – consentiva di analizzare i dispositivi meccanici come casi analoghi a creature organiche, ma prospettava soprattutto la possibilità di un incontro, di una relazione, di un'integrazione tra viventi – in particolare umani – e macchine.

(6) Cfr. Wiener N. (1948/1961), *Cybernetics: or control and communication in the animal and the machine*, Hermann & Cie, Paris, MIT Press, Cambridge Mass (trad.it. *La Cibernetica*, Mondadori, Milano, 1968).

(7) Galison P. (1994), *The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision*, in "Critical Inquiry", vol. 21, No. 1 (Autumn), 228-66, cit. p. 229.

(8) La *Stanford School* è una scuola di filosofia della scienza di impostazione pluralista, cioè contro l'idea della scienza come un tutto unitario, i cui membri originari erano passati per la Stanford University tra gli anni Ottanta e Novanta. Le tesi di questo gruppo di studiosi si concentravano sulla natura plurale dei modelli scientifici e su un approccio post-positivistico basato sulle pratiche della conoscenza scientifica.

(9) Cfr. Galison P. *The ontology of the enemy*, op. cit. p. 233.

(10) Rosenblueth A., Wiener N., Bigelow J. (1943), *Behavior, Purpose and Teleology*, in "Philosophy of Science", vol. 10 (1) (Jan), 18-24, (trad. it. *Comportamento, scopo, teleologia*, in Somenzi V., Cordeschi R., (a cura di) *La filosofia degli automi*, Bollati Boringhieri, Torino, 1994, 78-85).

Se le macchine potevano essere integrate negli esseri umani come una loro parte – come suggerivano le ricerche della cibernetica – il problema della radicale separazione tra loro sarebbe stato superato. Sarebbe stata, cioè, scavalcata la principale mancanza dell'intelligenza artificiale, quella legata alle capacità del corpo, a cui si sarebbe potuto sopperire attraverso una integrazione con il corpo dell'utente implicato nella macchina attraverso un'integrazione con lei.

Tale progetto di integrazione tra esseri umani e macchine e della conseguente loro collaborazione basata sul principio della comunicazione e del controllo fu molto proficuo, più di quello della tradizione dell'intelligenza artificiale almeno fino agli anni Duemila. Dall'interazione uomo-macchina derivarono alcuni dei maggiori successi dell'informatica: dal sistema operativo che si propone come *user-friendly*, basato sulla programmazione orientata agli oggetti (*object-oriented programming*), che introduce icone, finestre, *groupware*, e la capacità intuitiva di cliccare con il mouse o col dito sulle icone. Tutte queste innovazioni derivano dall'intuizione della possibilità di una simbiosi umano meccanica suggerita da Licklider (11). L'integrazione con la macchina era alternativa alla sostituzione meccanica dell'attività umana nello svolgere compiti intelligenti, prospettata dall'IA. L'integrazione, infatti, prevedeva una collaborazione nella quale ciascun agente avrebbe mantenuto le proprie caratteristiche più qualificanti integrandosi con l'altro:

To think in interaction with a computer in the same way that you think with a colleague whose competence supplements your own will require much tighter coupling between man and machine (12).

La simbiosi con la macchina derivava dalla cibernetica come disciplina che si occupava della comunicazione e del controllo. Il contributo rivoluzionario all'IA della cibernetica consiste nella riflessione sulla centralità della comunicazione tra esseri umani e macchine, e dalle conseguenze di questa impostazione derivino i maggiori successi della disciplina in generale, e l'entrata in una nuova agenda di ricerca. Comunemente invece si ritiene che, mentre l'IA fosse concentrata sulla riproduzione della mente, la cibernetica fosse interessata alla riproduzione del cervello. Sebbene questi due orientamenti distinti fossero in atto, erano solo modelli astratti che la macchina si incaricava di simulare senza avere molto a che fare né con la mente, né con il cervello umano. Erano solo diverse impostazioni su come costruire un modello efficiente dell'attività umana (13). L'aspetto della riproduzione dell'attività neurale sarà poi oggetto del progetto di ricerca del connessionismo (14), che spesso viene associato allo sviluppo delle macchine intelligenti della cibernetica. Eppure, l'articolazione della *human-machine interaction* è stata l'eredità della cibernetica che ha avuto il maggior impatto sull'IA, perché consentì di lanciare un metodo per la cattura dei modelli cognitivi umani necessari per l'evoluzione delle macchine e la loro progressiva autonomizzazione.

Fu Licklider, inoltre, che pure aveva partecipato alle fasi progettuali della cibernetica (insieme a Robert Taylor, a promuovere concretamente l'idea del computer come strumento di comunicazione (15)). I due sollevarono l'esigenza di un dispositivo in grado di connettere macchine differenti, capaci sia di interagire tra loro, sia di mediare con gli esseri umani. L'obiettivo delle macchine interconnesse era aperto a ogni possibile uso, anche quelli previsti o non ancora ipotizzati:

These new computer systems we are describing differ from other computer systems advertised with the same labels: interactive, time-sharing, multiaccess. They differ by having a greater degree of open-endedness, by rendering more services, and above all by providing facilities that foster a working sense of community among their users (16).

Tale articolo fu l'anticipazione della nascita di Arpanet, l'antesignana della rete Internet che si realizzò solo un anno dopo, sotto la guida di Taylor nell'ambito di un ufficio di ARPA (successivamente divenuta DARPA) chiamato Ipto (*Information processing technique office*).

A very important part of each man's interaction with his on-line community will be mediated by his OLIVER [...]. At your command, your OLIVER will take notes (or refrain from taking notes) on what you do, what you read, what you buy and where you buy it. It will know who your friends are, your mere acquaintances. It will know your value structure, who is prestigious in your eyes, for whom you will do what with what priority, and

(11) Licklider J.C.R. (1960), *Man-Computer Symbiosis*, in *IRE Transactions on Human Factors in Electronics*, vol. HFE-1, 4-11; consultato il 20 settembre 2024).

(12) Cfr. *ivi*, 4.

(13) Vedi Cordeschi, R., (2002), *The discovery of the artificial*, Springer Netherland, Dordrecht, per una descrizione precisa dei diversi modelli di intelligenza in gioco all'epoca: dall'emulazione della mente attraverso un sistema fisico simbolico al connessionismo che invece si proponeva di emulare una rete neurale in cui i singoli nodi erano unità elementari che agivano in parallelo con tutti gli altri nodi collegati da una rete che si attivava a seconda della quantità di stimoli ricevuti.

(14) Vedi *infra* per maggiori dettagli.

(15) Cfr. Licklider J.C.R., Taylor R.W. (1968), *The computer as a communication device*, in "Science and Technology: For the Technical Men in Management", No 76, April, 21-31.

(16) *Ivi*, 31.

who can have access to which of your personal files. It will know your organization's rules pertaining to proprietary information and the government's rules relating to security classification (17)

La lungimiranza di Licklider e Taylor non solo anticipò e indirizzò la nascita della rete Arpanet, ma si spinse anche a introdurre la possibilità di ottenere dati e correlazioni sui comportamenti delle persone ai fini di anticiparne e orientarne i comportamenti. OLIVER, che rappresenta una specie di assistente personale onnisciente che i due studiosi prevedono, è l'antesignano di tutta la nuova serie di servizi e progetti legati all'analisi dei Big Data nell'ambito della ricerca sociale, oltre a incarnare un primo esempio di assistente personale digitale, e anche in un certo senso dei vari dispositivi di Intelligenza artificiale generativi del tipo di ChatGPT.

5. Tra reasoning e learning: diverse strategie al confronto

Per entrare nel merito della distinzione tra *reasoning* e *learning* che ha costituito la storia dell'IA fin dalle origini dobbiamo partire da quel filone che si riprometteva di simulare le reti neurali del cervello e le loro capacità computazionali sub-simboliche. A questa tendenza apparteneva Frank Rosenblatt che nel 1957 costruì il Perceptron (1958), seguendo la tradizione di ricerca che aveva mosso i primi passi con l'articolo di McCulloch e Pitts del 1943. L'intuizione del Perceptrone era costruire una rete neurale in cui ogni nodo era una versione molto limitata di un neurone che apprendeva attraverso un processo di prova ed errore (Somenzi e Cordeschi 1994, passim).

Il sistema, tuttavia, non funzionava molto bene, sia perché era troppo elementare, sia per mancanza di dispositivi hardware sufficientemente potenti da sostenere la richiesta di capacità di calcolo di cui la macchina avrebbe avuto bisogno per i risultati. Il progetto venne interrotto da un famoso libro di grande influenza scritto da Minsky e Papert nel 1969, dal titolo *Perceptrons*. In questo libro si dimostrava che l'uso di una rete a due livelli come era quella di Rosenblatt non avrebbe garantito di apprendere come categorizzare degli aspetti che si proponeva di riconoscere. Nel libro si sosteneva che, sebbene fosse possibile che una rete a più livelli potesse ottenere dei risultati interessanti, non si poteva mai assicurare che sarebbe stata in grado di addestrarsi per garantire risultati soddisfacenti. Il successo del libro fu complice di un cambio di direzione delle ricerche del settore che abbandonarono le reti neurali. Trascorsi alcuni anni, però, l'interesse per questi strumenti aumentò. Da un lato si rianimò il connessionismo – per il quale ci fu un primo revival negli anni Ottanta con la pubblicazione di *Parallel distributed processing* (18) (19) – dall'altro si attivarono gli studi sul *machine learning*, sia pure in contesti minoritari e con scarsi finanziamenti (20).

Le ricerche prevalenti nell'ambito dell'IA tra la fine degli anni Sessanta e gli anni Ottanta erano invece, fondate sulla conoscenza simbolica (21). Nell'ambito del *machine learning*, si usava l'apprendimento basato sulla conoscenza, che prevedeva, in assenza della grande quantità di dati, ora disponibile, di costruire basi dati di informazioni da inserire nel dispositivo, magari utilizzando un linguaggio di programmazione di tipo dichiarativo. Tale strumento consentiva di trarre inferenze per nuove scoperte, partendo dalle conoscenze inserite in memoria. Questo sistema, però, era lento, rigido richiedeva un grande costo computazionale, eppure fino agli anni Novanta era ancora molto in voga nell'ambito dell'IA. Veniva a volte chiamato, in modo un po' dispregiativo GOF AI (Good Old Fashioned AI), cioè Intelligenza artificiale Classica o all'antica (22). Fino a che non erano disponibili i Big Data, però, non c'erano alternative, perché il sistema delle reti neurali non funzionava bene, a causa dell'assenza dei dati per addestrare le reti e di quelli per testare l'addestramento. Con la disponibilità delle banche dati, rese accessibili attraverso la rete, e con lo sviluppo delle GPU (graphic processing unit), che si rivelarono più adeguati dei normali processori a gestire grandi capacità di calcolo, l'interesse per il *learning* aumentò. Un manipolo di studiosi che non avevano mai smesso di credere alle reti neurali divenne improvvisamente *mainstream*. Si trattava di Geoff Hinton, Yoshua Bengio, Yann LeCun i quali, non a caso, nel 2018 hanno ricevuto insieme il Turing Award, considerato il Nobel per l'informatica (23).

(17) Ivi, 38-39.

(18) Rumelhart, D.E., Hinton, G.E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*, MIT Press, Cambridge, MA.

(19) Vale la pena segnalare che tra gli autori di questo volume c'è proprio Geoffrey Hinton a cui nell'ottobre del 2024 è stato conferito il premio Nobel per la fisica, rendendolo il primo studioso ad aver avuto sia il Nobel sia il Turing Award per l'informatica. Hinton ha inventato delle reti neurali artificiali sul modello di alcuni fenomeni della termodinamica, chiamate macchine di Helmholtz, oltre ad aver partecipato alla diffusione della tecnica della Backpropagation, uno degli algoritmi più utilizzati per addestrare una rete non supervisionata, ed essere a capo del progetto che nel 2012 costruì una rete non supervisionata Alexnet che aveva vinto la sfida di classificare le immagini correttamente promossa dalla banca dati ImageNet di cui parleremo più avanti.

(20) Cfr. Marcus G., Davis E. (2019), *Rebooting AI: building artificial intelligence we can trust*, Pantheon, New York, cap. 3.

(21) Cfr. Numerico T. (2021) *Big Data e Algoritmi*, Carocci, Roma, cap.1.

(22) Cfr. Boden, M.A. (2018). *Artificial intelligence: A very short introduction*. Oxford University Press, Oxford, Trad. it *Intelligenza Artificiale*, Il Mulino, Bologna, 2019.

(23) Cfr. Marcus, Davis Rebooting Ai, *op. cit.*, cap. 3, *passim*.

Uno dei motivi di grande successo del *machine learning* basato sulle reti neurali è che a fronte dei tanti dati disponibili per l'addestramento, la macchina non ha bisogno di troppo lavoro ingegneristico esperto perché viene poi affinata attraverso la grande quantità di tentativi che conduce per dare senso ai dati. Questo processo è molto diverso da quello in cui tutte le conoscenze dovevano essere inserite nella macchina sotto forma di sistemi esperti, come accadeva nel vecchio tipo di apprendimento in cui la base dati doveva essere esplicita.

Tuttavia, il risultato di Minsky e Papert non è mai stato confutato. Attraverso i meccanismi di apprendimento basati sulle reti neurali non possiamo essere sicuri dei risultati che otteniamo perché potrebbero non essere del tutto corretti, sebbene si possano ottenere risultati interessanti (24). Significa anche che se utilizzeremo questi sistemi di apprendimento in contesti delicati come cure mediche, o decisioni di giustizia sociale non potremmo rendere conto del perché di certi esiti e non potremmo essere sicuri che non contengano falsi positivi o falsi negativi, cioè situazioni nelle quali la valutazione è scorretta o perché considera la presenza di alcune caratteristiche nei soggetti sbagliati (falso positivo), o perché non si accorge che alcuni soggetti hanno una certa caratteristica (falso negativo).

6. Machine learning

La rilevanza delle tecniche del *learning*, e in particolare del *deep learning* hanno diverse ragioni, tra le quali spicca lo sviluppo di processori più performanti che in passato e la grande disponibilità dei corpora di addestramento, oltre agli algoritmi di apprendimento sviluppati nel corso di diversi decenni. Gli strumenti attuali sono gli eredi delle tecniche di reti neurali che simulavano il funzionamento cerebrale. Lo sviluppo degli algoritmi di apprendimento usati per addestrare i sistemi era già avvenuto negli anni zero del duemila, ma il 2012 costituisce un punto di svolta. Il sistema AlexNet (25) una rete di *deep learning*, vinse la famosa sfida per il riconoscimento di immagini (26), distaccando notevolmente il secondo classificato.

I sistemi di *machine learning* e specialmente quelli di *deep learning*, che hanno avuto così tanto successo hanno tre principali obiettivi: azioni di riconoscimento, attività di classificazione con l'obiettivo della previsione e del contestuale supporto alla presa di decisione o *automated decision making*, compiti generativi, come la produzione di contenuto testuale o multimediale generato a partire da un prompt di ricerca.

Si tratta di attività differenti che, però, hanno in comune alcune caratteristiche. Affinché questi sistemi abbiano successo è necessaria la disponibilità di grandi quantità di dati pertinenti all'ambito a cui il sistema si applica. Per il riconoscimento di volti o immagini è necessario disporre di un database di immagini adeguatamente etichettato, per il supporto alla presa di decisione è necessario un database di informazioni rilevanti sulle quali addestrare il sistema a estrarre le regolarità e proiettarle sul futuro dopo aver creato adeguati criteri per classificare i dati. E abbiamo bisogno di un grande *corpus* testuale e di immagini e di file multimediali per estrarre nuovi prodotti multimediali o i nuovi testi richiesti.

7. La centralità dei dati

Le diverse applicazioni, quindi, condividono la richiesta di una disponibilità adeguata di informazioni pertinenti sulle quali esercitare gli algoritmi per scegliere come classificarli per estrarne le caratteristiche utili al riconoscimento, alla decisione o alla produzione. Tali dati contribuiscono, talvolta in modo estremamente rilevante a orientare i risultati di questi sistemi, e sono parte integrante del processo cosiddetto 'intelligente'.

La prima questione critica di queste procedure riguarda la disponibilità dei dati. Sappiamo che la rete internet e i suoi servizi e specialmente l'avvento delle piattaforme hanno consentito la raccolta sistematica di grandi quantità di dati in formato digitale. Tale opportunità ha offerto ai possessori delle piattaforme, e a coloro che da loro si rifornivano, di costruire un nuovo spazio di appropriazione nel quale vigesse una completa assenza di regole di tutela. Di chi sono, infatti, i dati personali disponibili in formato digitale sulle piattaforme proprietarie? Nei contesti scientifici raccogliere dati è costoso, complicato e obbedisce a precise regole molto stringenti, ma i dati sperimentali garantiscono che chi li usa abbia anche contribuito a crearli, abbia seguito degli standard disciplinari per la loro raccolta e li metta a disposizione della collettività dei propri pari per ulteriori verifiche e per replicare gli esperimenti. I dati digitali non seguono queste regole perché la raccolta avviene considerando appropriabili i contenuti resi digitali e non si riferisce a nessuna pratica ufficiale per la costruzione di campioni effettivamente rappresentativi dei fenomeni che sarebbero oggetto della ricerca. Abbiamo quindi un ribaltamento delle regole per la raccolta dei dati e osserviamo un processo di appropriazione e di conseguente estrazione di valore senza

(24) *Ibidem*.

(25) Cfr. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. (2012) Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems* 25, pubblicato ufficialmente in *Communications of the ACM*, 60(6), 84-90, 2017. Si tratta di un articolo cruciale sia per il risultato dirompente nella sfida del riconoscimento di immagini lanciata nel concorso di ImageNet voluto dal Lin Fei Fei, sia perché si usava il sistema dei grafical processing unit, i processori grafici che diventeranno da allora in poi lo standard per il funzionamento dei sistemi di *deep learning*. Erano, cioè, due innovazioni in una gli algoritmi di convolutional neural network usati per il riconoscimento delle immagini e l'infrastruttura tecnica per il processamento dei grandi dati necessari allo scopo.

(26) Il nome della sfida è: *ImageNet Large Scale Visual Recognition Challenge*.

giustificare il tipo d'uso che viene effettuato, né rispetto a chi li ha prodotti, né in relazione al rispetto di standard scientifici (27).

Il ruolo centrale svolto dai dati per l'addestramento nel machine learning nell'esercitare le valutazioni impedisce la spiegabilità dei risultati ottenuti, soprattutto quando l'obiettivo del processo di apprendimento non può essere chiaramente esplicitato, come invece accade quando la funzione target di un processo algoritmico di addestramento riguarda la vittoria nei giochi come scacchi e go. Il sistema algoritmico del *deep learning* è fatto in modo da non esercitare un controllo diretto sugli strati interni della rete, conseguentemente la relazione tra processo di addestramento e suoi esiti si limita a considerare solo le prestazioni dell'output. Le procedure del *deep learning* non rispettano una delle caratteristiche definitorie degli algoritmi: la riconoscibilità del rapporto tra un passo e i suoi effetti sulla procedura.

Un altro aspetto su cui vale la pena riflettere criticamente è la relazione che intercorre tra conoscenza e creatività nell'uso degli strumenti digitali. Già Turing nel 1950 (28) affermò che, se vogliamo che le macchine abbiano comportamenti intelligenti, non possiamo impedire loro di commettere degli errori.

Nelle attività di riconoscimento la questione dell'errore si pone per il fatto che i processi algoritmici preposti a questo profilo, oltre a dipendere dall'accuratezza della base dati, funzionano come procedure di discriminazione che selezionano una certa forma discriminandola dalle altre in base alla differenza. Il processo umano di riconoscimento è diverso perché funziona per incremento di dettagli e non per selezione di forme a partire dalle differenze (29). Tale caratteristica implica necessariamente la possibilità di giudizi discriminatori che eliminano soggettività o aumentano gli errori in certi ambiti in cui le forme non si possono costituire per assenza di dettagli disponibili. È esemplare, a questo proposito, la difficoltà dei software di riconoscimento facciale nel distinguere le donne afroamericane, per le quali forse non ci sono abbastanza dettagli (30).

L'altro problema molto rilevante riguarda la capacità di questi sistemi di prevedere il futuro e intervenire a valutare o a manipolare le attività dei singoli soggetti di cui si detengono i dati come conseguenza della previsione.

Sappiamo che l'uso dei sistemi di supporto alla presa di decisione si basa sull'esercizio della statistica e sul principio di induzione che nulla possono dire sul caso singolo, e non possono garantire la verifica delle predizioni future, soprattutto se hanno per oggetto le persone singole, come nel caso in cui siano implicati meccanismi sociali come l'accesso al welfare o al credito, la definizione del premio assicurativo, la valutazione del rischio di pericolosità, il reclutamento del personale.

8. Criteri di validazione della conoscenza

Il futuro è incerto e, soprattutto rispetto ai singoli soggetti, nulla valgono le previsioni collettive, eppure tali previsioni esercitano un enorme potere a causa della loro provenienza da fonti accreditate che le trasformano in prescrizioni che cambiano la vita delle persone sulle quali vengono esercitati i giudizi (31).

Nella discussione tra conoscenza e creatività dobbiamo ammettere che questi sistemi non possano essere accreditati come strumenti conoscitivi, almeno rispetto all'idea che ne abbiamo in relazione ai criteri di validazione adottati a partire dallo sviluppo della scienza moderna. Nonostante l'instabilità e l'incertezza o, peggio, il potenziale discriminatorio di queste previsioni, tendiamo a pensare di poter delegare a questi sistemi di supporto alla presa di decisione la responsabilità di scelte dolorose e intrusive relative a diritti e equità delle persone che vivono nei contesti democratici. Oltre al problema che queste decisioni vengono prese da una scatola nera, cioè da un sistema algoritmico opaco e inaccessibile, abbiamo la questione della legittimità di proiettare il passato sul futuro nel caso di situazioni contingenti e imprevedibili, laddove l'addestramento basato sui dati concreti potrebbe naturalizzare differenze e marginalità che sono storicamente il frutto di passate ingiustizie.

Diverso sarebbe il discorso se questi strumenti di supporto decisionale fossero solo affiancati a decisori umani, costruendo percorsi di interazione esplicita e fornendo le chiavi di comprensione dei processi inferenziali che sono al lavoro nei sistemi algoritmici usati per l'apprendimento delle regolarità dei dati, detto *pattern recognition*.

9. Intelligenza artificiale generativa: la creatività dei modelli

Gli ultimi clamorosi risultati dell'intelligenza meccanica riguardano i cosiddetti *foundation models*, cioè quegli strumenti che, basandosi su grandi corpora di immagini, suoni e testi sono in grado di restituire contenuti creativi come output di un prompt di richiesta più o meno dettagliato. Il successo degli ultimi due anni è dovuto allo sviluppo di una tecnologia chiamata transformer che è il frutto di un articolo citatissimo dal titolo *Attention is all*

(27) Beaulieu A., Leonelli S. (2022), *Data and Society*, Sage, London.

(28) Turing A. 1950, *Intelligent Machinery*, op. cit.

(29) Chun, W.H.K. (2021). *Discriminating data: Correlation, neighborhoods, and the new politics of recognition*. MIT press, Cambridge (Mass.).

(30) Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification, in *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.

(31) Hildebrandt, M. (2021). *The issue of bias. The framing powers of machine learning* (dec. 2019). In Pelillo M., Scanturlo T. (eds.) *Machine we trust. Perspectives on dependable AI*, Mit Press, Cambridge (Mass.).

you need (32). In questo testo si sottolinea l'importanza dell'interazione con la richiesta, il prompt, a cui si deve dare una speciale attenzione.

In questo caso siamo di fronte a un dispositivo capace di invenzione, privo di velleità di tipo conoscitivo, che si nutre ancora una volta della disponibilità di contenuti digitali, come dimostra la scelta di tanti giornali come il New York Times o il Manifesto di negare l'accesso ai propri dati per l'addestramento. Il grande successo di questi strumenti come ChatGPT che ha avuto in pochi giorni più di 100 milioni di utenti nel novembre 2022 riposa anche sulle capacità mimetiche di questi sistemi che spesso finiscono per restituire direttamente parti dei corpora di addestramento. Il funzionamento di questi metodi per ottenere la creatività dipende dalla previsione statistica di una rosa delle risposte più frequenti in relazione al prompt di ricerca. Si tratta, quindi, di riprodurre un'aspettativa il più possibile prevedibile. Alcuni studiosi si spingono a ritenere che la produzione di testi e altri contenuti prefiguri l'attivazione di una vera autonomia soggettiva da parte del sistema (33), ma le frequenti allucinazioni, la mancanza di conoscenze di tipo comune, l'assenza di rapporto con il mondo esterno rendono il dispositivo produttivo, ma difficilmente capace di una vera intelligenza. In relazione a questa produzione di contenuti originali, entro certe definizioni di intelligenza, potrebbe risultare plausibile includere anche gli agenti sociotecnici implicati nella produzione di contenuti nella lista di quelli intelligenti, sebbene questo concetto non si possa precisare in modo univoco.

Questi strumenti aprono però nuovi scenari critici in quanto implicano l'appropriazione dei corpora e lo sfruttamento dell'esercito di persone coinvolte nella loro annotazione e nel processo di *Reinforcement learning by human feedback* (RLHF) che include il lavoro di alto livello di coloro che si esercitano a migliorare la qualità delle risposte fornite. Ma soprattutto attivano una nuova difficoltà per chi ne fruisce: quella di riconoscere se immagini, voci e descrizioni riguardano situazioni concrete di cui sono la testimonianza o introducono contenuti sintetici che nulla hanno a che fare con eventi e situazioni reali. Se non riusciremo a contenere quest'esplosione di ambivalenza, potremo trovarci nella situazione di non avere più una rappresentazione del mondo condivisa, sia pure criticamente, potremmo invece assistere all'esplosione di una molteplicità di rappresentazioni del mondo simili a quelle causate dalle teorie del complotto e dai terrapiattisti che per adesso si limitano a identificare minoranze prive di presa concreta sul mondo. Difficile dire come validare i contenuti per distinguere quelli immaginari e quelli oggetto della descrizione concreta del mondo, soprattutto per la potenziale esplosione quantitativa di immagini, suoni e testi, che risulterebbero impossibili da controllare attraverso la capacità cognitiva umana.

Possiamo anche riconoscere, nelle scelte inferenziali dei sistemi generativi, esiti discriminatori e diffusori di stereotipi. Osserviamo, come è segnalato in alcuni articoli del gruppo di Sasha Luccioni (34) di *Hugging Face*, che ad alcune richieste del *prompt*, relativamente alle immagini di persone che ricoprono certe professioni, corrisponda una rappresentazione pregiudiziale, etnicamente orientata che esibisce delle esplicite distorsioni della realtà a favore di standard, concentrati su una visione tradizionalista. La ricerca valuta il risultato di *Text to image*, cioè risposte a *prompt* testuali con immagini dei sistemi commerciali di IA generativa più noti, in relazione a diverse professioni rispetto a genere e appartenenza etnica delle persone che vengono riprodotte nelle immagini richieste, osservando il carattere profondamente discriminatorio e l'esclusione di identità marginalizzate in relazione a certe professioni, considerate più rispettabili o di successo. Tali risultati sono ancora più penalizzanti rispetto alle soggettività più fragili, tanto da superare, nella rappresentazione standard, la fenomenologia concreta della presenza reale etnica e di genere di professioni di alto livello come giudici, medici, manager ecc. La distorsione rappresentativa dei dati sintetici generati dalle richieste degli utenti amplifica la visione tradizionale del mondo polarizzandone i contorni.

Gli stessi progettisti dei sistemi generativi devono segnalare il carattere allucinatorio di alcune risposte, proponendosi di migliorare le prestazioni con le versioni successive dei *software* (35). Ha colpito la fantasia collettiva l'*exploit* del *chatbot* di Google, Gemini, che suggeriva di aggiungere la colla alla salsa della pizza (36), tuttavia, usare il termine "allucinazioni" potrebbe essere fuorviante. Il significato di questo termine, infatti, si riferisce a errori di carattere precettivo dei dispositivi che potrebbero contribuire ad amplificare la retorica della quasi umanità dei sistemi generativi. Forse varrebbe la pena chiamare i risultati erronei con il loro nome come provano a

(32) Vaswani, A. Shazeer, N. Parmar, N. U. et al *Attention is all you need*, Advances in Neural Information Processing Systems. 30. Curran Associates, Inc.

(33) Cfr. Roncaglia G. (2023) *L'architetto e l'oracolo*, Laterza, Roma-Bari.

(34) Cfr. Luccioni, A.S., Akiki, C., Mitchell, M., & Jernite, Y. (2023). Stable bias: Analyzing societal representations in diffusion models. arXiv preprint arXiv:2303.11408. Per un articolo più divulgativo sull'argomento dallo stesso gruppo di ricerca, rimando al blog dell'azienda francese Hugging face, per la quale lavora Sasha Luccioni, che è la principal investigator di questo progetto su IA generativa e pregiudizi strutturali, questo il post sulla newsletter aziendale: Sasha Luccioni et al. (26/06/2023) Ethics and Society Newsletter #4: Bias in Text-to-Image Models.

(35) Cfr. Prabhakar Raghavan Gemini image generation got it wrong. We'll do better, Google Blog 23/02/2024, ma secondo i commentatori ancora a maggio 2024 le cose per il generatore di immagini Gemini di Google non sono migliorate, vedi Kyle Wiggers, *Google still hasn't fixed Gemini's biased image generator*, TechCrunch, 15/05/2024, e più in generale Alex Cran, *We have to stop ignoring AI's hallucination problem*, The Verge, 15/05/2024.

(36) Cfr. Elizabeth Lopatto, *Google still recommends glue for your pizza*, The Verge, 12/06/2024.

fare un gruppo di studiosi che titolano il loro articolo *ChatGPT is bullshit*. Secondo questo lavoro – ma non sono gli unici o i primi a notarlo – attribuire all’IA generativa delle allucinazioni è mistificatorio e scorretto. “Riteniamo che queste falsità, la generale attività dei *large language models*, si possa comprendere meglio come *bullshit* [...] i modelli sono, in modo rilevante, indifferenti alla verità dei loro risultati” (37). Varrebbe la pena riflettere sul perché tutto questo clamore per dei sistemi non interessati o impossibilitati a rispondere correttamente alle nostre domande, che richiedono una costante e attenta capacità di *editing* e di correzione attiva, oltre che orientati dai dati e dai vincoli loro imposti a veicolare una rappresentazione della società e delle sue relazioni che restituisce le preferenze dei più forti, amplificandole rispetto alle soggettività e alle loro condizioni di privilegio o di marginalità.

10. La politica dei sistemi intelligenti

Le trasformazioni promosse dall’uso degli strumenti di intelligenza artificiale hanno un carattere fortemente politico. La scelta di delegare decisioni, soprattutto quelle che riguardano il futuro, come nel caso della giustizia predittiva non può essere risolto da meccanismi di automazione. Ci troviamo in contesti incerti e contingenti. Cedere il passo a oracoli instabili e incomprensibili non riguarda tanto una loro presunta superiorità, ma un eventuale spinta, più o meno consapevole, a evitare di assumere la responsabilità delle scelte da esercitare in contesti imprevedibili.

Un aspetto cruciale da considerare è la rilevanza dell’addestramento sui dati disponibili che contengono le disparità e le iniquità che sono attualmente vigenti nel mondo, per questo avere sistemi che inferiscono proiezioni future a partire da questi non è troppo rassicurante come mostrano molti studiosi (38).

Quanto alla volontà di affidarsi a sistemi generativi per la produzione dei contenuti, porterà molti cambiamenti nelle attività intellettuali, ma non è detto che rappresenterà un miglioramento delle condizioni di vita delle persone. Il processo di raccolta dei dati personali di miliardi di persone costituisce la condizione di possibilità di questi strumenti, ma si tratta di meccanismi di appropriazione di risorse comuni. Verrà un momento in cui questo passaggio incongruo diverrà chiaro per tutti. La possibilità di uso di certi strumenti non dà di per sé il diritto di esercitarla. Ci sono regolazioni che impediscono di usare i dispositivi alla loro massima potenza, come il codice stradale, o le regole sul porto d’armi. Sappiamo bene che non basta che una cosa sia possibile per essere autorizzata. Sebbene il contesto della raccolta dei dati sia complesso per il fatto che l’appropriazione si svolge in luoghi lontani dalla giurisdizione dove valgono le conseguenze legate all’estrazione di valore da quei contenuti.

Come suggeriscono acutamente Acemoglu e Johnson (39) non tutte le trasformazioni tecnologiche portano con sé un progresso per la collettività. Dobbiamo compiere delle scelte politiche ed epistemologiche per orientare queste innovazioni affinché possano migliorare le condizioni di vita degli umani e in generale dei viventi.

* * *

(37) Hicks, M.T., Humphries, J. & Slater, J. ChatGPT is bullshit. *Ethics Inf Technol* 26, 38 (2024).

(38) Qui ci limitiamo a citare alcuni testi estremamente espliciti a questo riguardo, Couldry N., Mejias U.A. (2019), *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*, Stanford University Press, Stanford (ca), trad.it. *Il Prezzo della connessione*, Il Mulino, Bologna, 2022, D’ignazio, C., & Klein, L.F. (2023). *Data feminism*. MIT press, Cambridge, Crawford K. (2021) *Atlas of AI*, Yale University Press, New Haven e London, Trad.it. *Nè intelligente, né artificiale*, Il Mulino, Bologna, 2021, Osoba O.A., Boudreaux B., Saunders J. et al. (2019), *Algorithmic Equity: A Framework for Social Applications*, RAND Corporation, Santa Monica, CA.

(39) Acemoglu D., Johnson S. (2023) *Power and progress*, *PublicAffairs*, New York, Trad. it. *Potere e progresso*, Il Saggiatore, Milano, 2023.

AI E FUTURO OCCUPAZIONALE: SFIDE E OPPORTUNITÀ IN UN MONDO IN EVOLUZIONE

di Lucia de Grimani (*)

Abstract: In un'epoca di rapida innovazione tecnologica, l'intelligenza artificiale (AI) sta trasformando profondamente il nostro modo di vivere e lavorare. Questo articolo esplora l'impatto dell'AI sul mercato del lavoro, evidenziando sia le opportunità che le sfide etiche e sociali che ne derivano. Attraverso l'analisi delle dinamiche di automazione e della necessità di aggiornamento continuo delle competenze, si sottolinea come l'AI non sia solo una tecnologia dirompente, ma anche uno strumento potenzialmente emancipatore. Tuttavia, per evitare che diventi fonte di alienazione e disuguaglianza, è essenziale adottare un approccio consapevole e trasparente. L'articolo discute l'importanza di politiche inclusive e di una governance responsabile per garantire che i benefici dell'AI siano equamente distribuiti e che questa diventi un motore di crescita sostenibile e giustizia sociale.

In an era of rapid technological innovation, artificial intelligence (AI) is profoundly transforming the way we live and work. This article explores the impact of AI on the labor market, highlighting both the opportunities and the ethical and social challenges it brings. Through the analysis of automation dynamics and the need for continuous skills development, it emphasizes how AI is not just a disruptive technology but also a potentially empowering tool. However, to prevent it from becoming a source of alienation and inequality, it is essential to adopt a conscious and transparent approach. The article discusses the importance of inclusive policies and responsible governance to ensure that the benefits of AI are fairly distributed and that it becomes a driver of sustainable growth and social justice.

Sommario: 1. Introduzione. – 2. AI e trasformazione del mercato del lavoro. – 3. Il ruolo della formazione e della riqualificazione professionale. – 4. AI, innovazione e sostenibilità. – 5. Le politiche aziendali e l'integrazione dell'AI – 6. AI: strategie per una transizione inclusiva. – 7. Umanesimo digitale: l'AI al servizio dell'umanità. – 8. AI come diritto universale.

1. Introduzione

Dopo il Rinascimento stiamo vivendo nel periodo a più alto tasso di innovazione di tutta la storia dell'umanità. Siamo nel mezzo di una rivoluzione tecnologica senza precedenti, in cui il progresso e la scienza offrono opportunità mai viste prima. Questa è sicuramente la fase più rivoluzionaria nel percorso del genere umano, un'era che sarà ricordata come un punto di svolta per lo sviluppo di tecnologie che renderanno possibile, l'impossibile.

In questo scenario, l'intelligenza artificiale (AI) sta ridisegnando ogni aspetto della nostra vita, trasformando le modalità con cui lavoriamo, comunichiamo e persino comprendiamo il mondo che ci circonda. Se da un lato l'AI ci offre strumenti straordinari per automatizzare compiti ripetitivi e ampliare le nostre capacità cognitive, dall'altro ci obbliga a riflettere profondamente su ciò che significa essere umani in un'epoca in cui le macchine assumono un ruolo sempre più centrale nelle nostre vite.

Considerando questi aspetti, la metafora della Caverna di Platone si rivela particolarmente pertinente. Nell'allegoria, gli uomini, prigionieri di una caverna, vedono soltanto le ombre degli oggetti proiettate su una parete, confondendo quelle ombre per la realtà stessa. In maniera analoga, l'intelligenza artificiale, con la sua capacità di raccogliere, filtrare e interpretare enormi quantità di dati, rischia di farci vivere in una sorta di 'caverna digitale'. Le informazioni che ci vengono presentate, mediate da algoritmi e intelligenze artificiali, potrebbero non rappresentare la realtà oggettiva, ma una sua interpretazione parziale, influenzata da logiche di programmazione, bias cognitivi e obiettivi aziendali.

Così come i prigionieri nella caverna erano inconsapevoli della vera natura di ciò che li circondava, rischiamo oggi di essere intrappolati in una visione limitata e artificiale del mondo, modellata dagli strumenti tecnologici che utilizziamo. Tuttavia, se Platone indicava la filosofia come mezzo per liberarsi da questa illusione e accedere alla verità, oggi il nostro compito è garantire che l'intelligenza artificiale non diventi un ulteriore strumento di alienazione, ma piuttosto un mezzo per ampliare la nostra conoscenza e comprensione. La sfida non è solo tecnologica, ma epistemologica: dobbiamo essere in grado di riconoscere le 'ombre' generate dagli algoritmi e sviluppare una capacità critica che ci permetta di discernere la realtà dalle sue rappresentazioni digitali. In questo senso, l'AI può diventare sia un potenziale strumento di emancipazione, che ci libera dai compiti ripetitivi e ci consente di concentrarci su attività più creative e analitiche, sia una fonte di nuove forme di dipendenza intellettuale e manipolazione della percezione. Per evitare che ciò accada, è essenziale promuovere un uso consapevole e

(*) L. de Grimani è presidente CNA Impresa Donna.

trasparente della tecnologia, in cui i processi decisionali e le dinamiche algoritmiche siano chiaramente comprese e gestite, restituendo all'essere umano il controllo e la capacità di uscire 'dalla caverna' (1).

"AI is not another utility that needs to be regulated once it is mature. It is a powerful force, a new form of smart agency, which is already reshaping our lives, our interactions, and our environments" (2).

Mentre l'intelligenza artificiale ci apre a opportunità senza precedenti è fondamentale interrogarci sulla nostra reale preparazione a gestire con discernimento questo nuovo potere. Siamo pronti a navigare le complessità etiche e sociali che ne derivano, e quali misure possiamo adottare per assicurarci che tali strumenti siano utilizzati in modo responsabile e vantaggioso per l'intera società? Da sempre, l'uomo ha cercato di plasmare la realtà a sua immagine, ma mai come ora abbiamo la possibilità di farlo su scala così vasta e con un impatto così profondo.

L'ascesa dell'AI porta con sé una contraddizione fondamentale: da un lato, ci offre strumenti potenti per migliorare la nostra condizione, dall'altro, rischia di sottrarre all'uomo il controllo sulle sue stesse creazioni, spingendoci verso una realtà in cui la macchina diventa protagonista. Il pericolo non risiede nella tecnologia in sé, ma nel modo in cui noi esseri umani decidiamo di utilizzarla. La sfida più grande sarà garantire che l'intelligenza artificiale rimanga uno strumento di emancipazione collettiva, accessibile e utile a tutti, piuttosto che un privilegio di pochi.

Con l'obiettivo di fornire una visione completa, in queste pagine esamineremo, non solo come l'AI stia trasformando il mercato del lavoro, ma anche le implicazioni etiche, sociali e politiche che derivano dalla sua rapida adozione, perché, in ultima analisi, il vero progresso non è solo tecnologico: è quello che riguarda l'intera umanità.

2. AI e trasformazione del mercato del lavoro

L'intelligenza artificiale sta determinando una significativa discontinuità nel mercato del lavoro, specialmente in quei settori che basano il loro core business su mansioni ripetitive e standardizzate. *"The adoption of AI presents an opportunity to enhance productivity and reduce costs, but also introduces significant challenges, especially regarding the protection of workers at risk of job displacement due to automation. Affirming the principle of autonomy in the context of AI means striking a balance between the decision-making power we retain for ourselves and that which we delegate to artificial agents"* (3). Secondo il Fondo Monetario Internazionale, fino al 40% dei posti di lavoro globali potrebbe essere impattato dall'AI, con un'incidenza che sale fino al 60% nelle economie più avanzate. Questa rivoluzione è già in atto: grandi aziende, come IBM, stanno adattando le proprie strategie di assunzione per ridurre i ruoli che possono essere sostituiti dall'automazione, mentre altre realtà aziendali ripensano la propria organizzazione per poter integrare l'AI nei processi produttivi.

Il settore IT è stato tra i primi a sentire gli effetti della trasformazione in atto, ma anche industrie tradizionalmente meno digitalizzate, come il customer service e i servizi di traduzione, stanno subendo il fascino della novità e riducendo il bisogno di lavoro umano a favore di tecnologie AI più efficienti. Tuttavia, l'AI non si limita a cancellare posti di lavoro, ne crea di nuovi, soprattutto in ambiti che richiedono competenze tecniche avanzate, come lo sviluppo di algoritmi o la gestione dei dati.

Il lavoro freelance è un settore particolarmente esposto a questa trasformazione, attività come la scrittura, la programmazione o la creazione di contenuti digitali vengono automatizzate, con conseguenti riduzioni di opportunità e una pressione sui compensi. Studi recenti su piattaforme di lavoro autonomo come *Upwork* e *Fiverr* mostrano che la domanda per questi servizi è in declino, mentre cresce il bisogno di figure in grado di perfezionare i contenuti generati dalle macchine. Possiamo quindi affermare che l'intelligenza artificiale stia contribuendo a modificare la natura stessa delle mansioni svolte dagli esseri umani, il ruolo dei lavoratori, infatti, si sta spostando verso attività più cognitive e meno fisiche, in un processo che richiede competenze sempre più specializzate. Se da un lato le macchine sono in grado di elaborare enormi volumi di dati e prendere decisioni su base statistica, dall'altro resta insostituibile la capacità umana di valutare, interpretare e prendere decisioni basate su etica e contesto, su istinto, sui "sento che..." e "mal di pancia".

L'adozione crescente dell'AI, come sottolinea il prof. Luciano Floridi nel *"AI4People - An ethical framework for a good AI society: opportunities, risks, principles, and recommendations"*, rappresenta un'opportunità per aumentare la produttività e ridurre i costi. Tuttavia, questa evoluzione porta con sé sfide significative, in particolare per quanto riguarda la protezione dei lavoratori le cui mansioni sono a rischio di automazione. *Per affrontare queste sfide, è essenziale comprendere il contesto in cui si sviluppa il progresso tecnologico, che si muove secondo tre dinamiche principali: esponenziale, combinatoria e ricorsiva.* L'avanzamento esponenziale della tecnologia significa che le innovazioni non crescono più a un ritmo lineare, ma accelerano rapidamente, con effetti imprevedibili sul mercato del lavoro e sulle nostre società. A ciò si aggiunge il carattere combinatorio delle nuove tecnologie come intelligenza artificiale, IoT-Internet delle cose e deep learning, etc., che non operano in modo isolato, ma si intrecciano tra loro, potenziandosi a vicenda e aprendo nuove possibilità applicative. Infine, la natura

(1) Platone, *La Repubblica*, libro VII, trad. di Giovanni Reale, Milano, Bompiani, 2003, pp. 514-520.

(2) L. Floridi et al., *"AI4People-An ethical framework for a good AI society: opportunities, risks, principles, and recommendations"*, Ssrn electronic journal, dicembre 2018, p. 2, in <<https://ssrn.com/abstract=3284141>>.

(3) L. Floridi et al., *op. cit.*, p. 5.

ricorsiva di alcune di queste tecnologie, come l'AI stessa, permette loro di migliorare autonomamente, generando un ciclo continuo di sviluppo che amplifica ulteriormente il loro impatto. Questi tre elementi, uniti, stanno ridisegnando il panorama lavorativo globale, creando sfide e opportunità senza precedenti. Per sfruttare appieno i benefici dell'AI senza sacrificare l'occupazione, è necessaria una strategia di lungo termine che comprenda sia politiche di formazione continua, sia una visione dell'intelligenza artificiale come strumento di supporto e non come sostituto del lavoro umano. “*The opportunities AI offers for promoting human dignity and flourishing are significant, but they must be weighed against the risks of malicious manipulation or inadvertent overuse*” (4). La chiave sta nel saper integrare la tecnologia nei processi produttivi senza rinunciare alla centralità del contributo umano, promuovendo un equilibrio tra efficienza e valore del lavoro.

3. Il ruolo della formazione e della riqualificazione professionale

Nell'era dell'intelligenza artificiale, la formazione non può più essere concepita come un processo confinato tra le mura scolastiche. Il mondo del lavoro si evolve a un ritmo talmente rapido che ciò che impariamo oggi può diventare obsoleto domani. Per questo, l'apprendimento deve diventare un percorso continuo, che accompagni i lavoratori durante tutta la loro vita professionale, consentendo loro di adattarsi a un contesto in costante trasformazione. Le sfide poste dall'AI che automatizza mansioni sempre più complesse, richiedono una preparazione non solo tecnica, ma anche flessibile ed efficace, capace di rispondere alle esigenze di ruoli che cambiano.

In questo panorama emergono due concetti chiave: *upskilling* e *reskilling*. L'*upskilling* rappresenta l'acquisizione di competenze avanzate che migliorano quelle già possedute, mentre il *reskilling* si concentra sull'apprendimento di abilità completamente nuove per prepararsi a ruoli differenti. Questi due processi sono oggi fondamentali per chiunque voglia mantenere la propria competitività in un ecosistema lavorativo sempre più dinamico: le tecnologie cambiano, ma chi è in grado di aggiornarsi o reinventarsi può rimanere al passo, continuando a creare valore e trovare nuove opportunità, invece di essere superato dai cambiamenti.

Le università e gli istituti di formazione sono chiamati a rispondere a questa trasformazione, adattando i loro programmi per preparare i professionisti del futuro. “*Support the creation of educational curricula and public awareness activities around the societal, legal, and ethical impact of Artificial Intelligence*” (5). Non si tratta più solo di formare tecnici, ma di sviluppare talenti che sappiano interagire con le tecnologie in modo consapevole e strategico. I programmi accademici oggi dovrebbero integrare competenze digitali come l'intelligenza artificiale, l'analisi dei dati e la programmazione, ma anche abbracciare le *soft skills*. Il pensiero critico, l'immaginazione e la capacità di risolvere problemi complessi sono ciò che permetterà agli esseri umani di mantenere un ruolo centrale in un mondo sempre più automatizzato. Queste abilità umane, infatti, sono insostituibili, poiché attingono a ciò che le macchine non possono replicare: intuizione, empatia e giudizio etico.

In un mondo in cui l'intelligenza artificiale esegue compiti ripetitivi e meccanici, il vero potenziale umano risiede nella capacità di affrontare sfide creative, di collegare concetti in modi inaspettati, di innovare. L'AI non deve essere vista come un'alternativa alle capacità umane, ma come un potenziatore, un abilitatore, ovvero: uno strumento che permette alle persone di liberarsi da attività alienanti per concentrarsi su ciò che davvero conta. In questo senso, le *soft skills* non sono accessorie, ma centrali per garantire il successo professionale in un ambiente lavorativo che sta subendo una nuova modellazione grazie, o a causa, dell'intelligenza artificiale.

Ecco perché possiamo affermare che il futuro del lavoro non sarà segnato dalla sostituzione delle persone, ma dalla loro capacità di collaborare con la tecnologia per ottenere di più, meglio e in meno tempo. Questa è la vera sfida della formazione e della riqualificazione professionale: preparare le persone a non temere l'intelligenza artificiale, ma a usarla per amplificare il loro potenziale e contribuire in modo significativo a un'economia in continua evoluzione.

4. AI, innovazione e sostenibilità

L'intelligenza artificiale rappresenta una leva straordinaria per promuovere un'innovazione che non solo accelera la produttività, ma apre anche la strada a uno sviluppo più sostenibile. In questa prospettiva, il concetto di *twin transition* – la doppia transizione, digitale e green – offre un'opportunità unica per ridefinire il modo in cui le aziende operano e interagiscono con l'ambiente. L'integrazione di soluzioni basate sull'AI può migliorare l'efficienza dei processi produttivi, ridurre il consumo energetico, minimizzare gli sprechi e ottimizzare l'uso delle risorse naturali, tracciando così un percorso concreto verso la sostenibilità. Secondo il World Economic Forum, le soluzioni digitali potrebbero ridurre le emissioni globali fino al 20%, dimostrando il potenziale della digitalizzazione per affrontare la crisi climatica (6).

(4) L. Floridi et al., op. cit., p. 3.

(5) L. Floridi et al., op. cit., p. 21.

(6) Unlocking Digital Innovation for Net Zero>World Economic Forum Annual Meeting|World Economic Forum, in <<https://www.weforum.org/events/world-economic-forum-annual-meeting-2022/sessions/unlocking-digital-innovation-for-net-zero>>.

Un esempio lampante di questo potenziale è l'applicazione dell'AI nella gestione del ciclo di vita dei prodotti: dall'estrazione delle materie prime fino al riciclo e al riuso, ogni fase del processo può essere monitorata e ottimizzata. Questo non solo consente una riduzione degli sprechi lungo tutta la filiera, ma apre la strada all'adozione di modelli di economia circolare, in cui le risorse vengono continuamente recuperate e reintegrate nei processi produttivi. In questo modo, si minimizzano i rifiuti e si massimizzano i benefici economici e ambientali.

Tuttavia, la vera trasformazione richiede più della semplice adozione di nuove tecnologie. Per sfruttare appieno le potenzialità dell'AI e della *twin transition*, è essenziale che le persone e le organizzazioni siano pronte ad accogliere il cambiamento. *Non è sufficiente una trasformazione tecnologica: serve anche una trasformazione culturale.* Oggi, l'ostacolo principale non è la mancanza di risorse o “τέχνη”, ma una diffusa resistenza al cambiamento, alimentata da timori legati all'incertezza, dall'attaccamento a pratiche consolidate e da una limitata comprensione dei benefici delle nuove tecnologie.

A fronte di questa situazione, emerge la necessità di un approccio olistico, in cui la tecnologia e il capitale umano si rafforzino a vicenda. L'AI, da sola, non è sufficiente a garantire una transizione sostenibile: senza un cambiamento di *mindset* e una preparazione adeguata, molte aziende rischiano di non sfruttare appieno il potenziale delle innovazioni digitali. Inoltre, è cruciale bilanciare il progresso tecnologico con la responsabilità sociale, se non adeguatamente gestita, la trasformazione digitale può accentuare le disuguaglianze, concentrando i benefici in mano a pochi, mentre le comunità più vulnerabili rischiano di essere escluse da questo progresso.

Per garantire che l'AI contribuisca a un benessere diffuso, le aziende devono adottare un approccio etico e inclusivo. Come sottolinea il World Economic Forum, “*AI should promote justice and seek to eliminate all types of discrimination, ensuring shared benefit and prosperity*” (7) pertanto, è fondamentale che l'adozione dell'AI non solo migliori i processi produttivi, ma crei anche un impatto positivo per le persone, favorendo un accesso equo alle opportunità generate da queste innovazioni. Questo è particolarmente rilevante quando si affrontano questioni di giustizia sociale, come la disuguaglianza di genere. Per esempio, l'integrazione di piattaforme di intelligenza artificiale nel settore educativo rappresenta un'opportunità straordinaria per promuovere l'equità di genere. Attraverso sistemi di apprendimento personalizzati e programmi di *mentorship* automatizzati, le donne possono finalmente acquisire competenze preziose, preparandosi per carriere in settori storicamente dominati dagli uomini. Dare ad ogni individuo la possibilità di esprimere il proprio potenziale e talento è quanto di più giusto ed etico possa essere fatto. In parallelo, l'adozione di tecnologie che promuovono la flessibilità lavorativa, come il lavoro agile e le piattaforme di collaborazione online, offre alle persone l'opportunità di bilanciare in modo armonioso le proprie responsabilità professionali e familiari. Immaginiamo una madre che, grazie allo *smart working*, può gestire il proprio lavoro comodamente da casa, dedicando tempo ai propri figli senza compromettere la sua carriera. Queste soluzioni non solo eliminano le barriere che ostacolano l'accesso delle donne al mercato del lavoro, ma creano anche condizioni favorevoli affinché possano prosperare in ogni aspetto della loro vita (8). Per tutte queste ragioni, investire in iniziative tecnologiche che affrontano la disuguaglianza di genere va oltre la semplice tutela dei diritti: rappresenta un'opportunità per arricchire la vita di tutti. Creando un futuro in cui ogni persona, indipendentemente dall'identità, possa svilupparsi pienamente, contribuiamo a costruire una società più equa e sostenibile anche per le generazioni a venire.

5. Le politiche aziendali e l'integrazione dell'AI

L'integrazione dell'intelligenza artificiale nei processi aziendali non può essere vista solo come un miglioramento tecnologico; richiede una governance attenta e una visione strategica di lungo termine. Questo perché l'AI ha il potenziale di influenzare profondamente la struttura aziendale, non solo in termini di produttività, ma anche di organizzazione del lavoro e impatti sociali. Le aziende devono considerare non solo l'efficienza, ma anche il benessere dei loro lavoratori e delle comunità coinvolte, adottando un approccio che favorisca la crescita delle competenze e la riqualificazione professionale.

Le aziende di successo bilanciano i vantaggi in termini di produttività con politiche che tutelano i lavoratori, promuovendo la loro crescita professionale e facendo sì che l'automazione non rappresenti una minaccia, ma un'opportunità. Questo può essere raggiunto coinvolgendo tutte le parti interessate, sia interne che esterne, in un dialogo costante. “*Creating a Good AI Society requires a multistakeholder approach, ensuring that AI will serve the needs of society, enabling developers, users, and rule-makers to all collaborate from the outset*” (9). L'inclusione dei lavoratori, la comunicazione trasparente e la formazione continua permettono di affrontare i timori legati alla perdita di posti di lavoro, trasformando l'automazione in un fattore di crescita anziché di esclusione.

Promuovere un dialogo aperto è fondamentale per ridurre le preoccupazioni dei lavoratori, mostrando loro come l'AI possa migliorare il loro lavoro anziché sostituirlo. Le aziende che abbracciano un approccio inclusivo e visionario dimostrano come l'adozione dell'AI possa andare oltre l'efficienza operativa, trasformandosi in uno

(7) L. Floridi *et al.*, *op. cit.*, p. 13.

(8) Women's Empowerment Observatory/The European House-Ambrosetti, in <<https://www.ambrosetti.eu/en/women-empowerment-observatory/>>.

(9) L. Floridi *et al.*, *op. cit.*, p. 16.

strumento per rafforzare il benessere dei dipendenti e favorire una sostenibilità sociale concreta. L'obiettivo non è solo incrementare la produttività, ma dare spazio a un'occupazione che valorizzi il contributo umano, facendo emergere talenti, idee e prospettive che solo la presenza attiva e consapevole di ogni persona può apportare. In questo modo, l'IA diventa uno strumento di inclusione e partecipazione, aprendo la strada a un nuovo modello di impresa, più capace di adattarsi alle esigenze della società e di contribuire a un cambiamento positivo per tutti.

6. AI: strategie per una transizione inclusiva

Affrontare la trasformazione indotta dall'intelligenza artificiale in modo sostenibile richiede un approccio inclusivo, che coinvolga attivamente sia le istituzioni pubbliche che il settore privato. Questo approccio è cruciale perché solo una collaborazione sinergica può garantire che tutti gli attori – governi, aziende e lavoratori – siano preparati e supportati durante la transizione. Le politiche del lavoro devono mettere al centro il lifelong learning e la riqualificazione professionale, con un'attenzione particolare allo sviluppo delle competenze chiave necessarie per operare in un contesto tecnologico avanzato, come la padronanza delle tecnologie digitali, l'analisi dei dati e la capacità di lavorare efficacemente con sistemi automatizzati.

Gli incentivi pubblici sono strumenti fondamentali per facilitare la transizione, così come le agevolazioni fiscali per le aziende che investono in programmi di formazione per i dipendenti, i programmi di supporto e i corsi di riqualificazione finanziati dal governo e dalle istituzioni in genere. Queste misure possono aiutare i lavoratori che rischiano di essere penalizzati dall'automazione, offrendo loro la possibilità di aggiornare le proprie competenze o di acquisirne di nuove, per restare competitivi. Le aziende, dal canto loro, possono adottare modelli di business inclusivi e sostenibili, che non puntino solo alla massimizzazione dei profitti, ma che integrino anche l'innovazione sociale. Un esempio concreto è l'implementazione di modelli di collaborazione uomo-macchina, in cui l'intelligenza artificiale non sostituisce il lavoratore, ma ne amplifica le capacità, consentendo ai dipendenti di concentrarsi su compiti più strategici e creativi, mentre l'AI si occupa delle attività ripetitive, frustranti o amministrative. Questo tipo di sinergia permette di ottimizzare la produttività e al tempo stesso migliorare le condizioni di lavoro, valorizzando il contributo umano.

L'utilizzo di tecnologie AI può diventare un motore di crescita sostenibile solo se viene integrata con un approccio che promuova anche il benessere sociale, in un mondo che pone sempre più al centro questi temi, le aziende che riescono a bilanciare l'innovazione tecnologica con una governance responsabile avranno un vantaggio competitivo. Oltre agli interventi pratici per un uso inclusivo dell'intelligenza artificiale è fondamentale non dimenticare di adottare una visione più ampia che promuova un *umanesimo digitale* e orienti la tecnologia al servizio degli individui, questo approccio non solo sostiene politiche responsabili, ma garantisce che il progresso tecnologico metta sempre l'essere umano al centro.

7. Umanesimo digitale: l'AI al servizio dell'umanità

L'umanesimo digitale propone una visione della tecnologia come uno strumento per valorizzare l'essere umano, piuttosto che sostituirlo o limitarlo. Si fonda sull'idea che l'essenza dell'umanità risieda in qualità distintive come la creatività, l'empatia, il giudizio morale e la capacità di adattarsi. Di conseguenza, l'intelligenza artificiale (AI) dovrebbe essere sviluppata con l'obiettivo di rafforzare queste caratteristiche, potenziando ciò che ci rende unici, senza tentare di replicarle o imitarle.

Questa prospettiva rifiuta sia la riduzione dell'uomo a un semplice insieme di algoritmi, sia l'idea che le macchine possano raggiungere una forma di intelligenza equivalente a quella umana (10). L'umanesimo digitale sottolinea le profonde differenze tra la mente umana e l'AI, considerando la tecnologia un valido sostegno per superare le limitazioni fisiche e cognitive, mantenendo però la centralità dell'essere umano.

L'intelligenza artificiale, quindi, assume il ruolo di strumento per ampliare le opportunità di apprendimento, innovazione e miglioramento delle condizioni di vita, a condizione che rimanga sotto il controllo umano. La tecnologia deve sempre servire i bisogni umani, non il contrario. Questo è il cuore dell'umanesimo digitale: una filosofia che rispetta l'autonomia della persona, ponendola come guida consapevole nello sviluppo e nell'uso della tecnologia.

Il vero pericolo, quindi, non risiede nella tecnologia in sé, ma nel modo in cui la utilizziamo. Se l'AI viene impiegata in modo etico e inclusivo, può diventare uno strumento potente per il bene comune. Tuttavia, un uso non regolamentato e privo di una riflessione sulle sue implicazioni etiche rischia di creare una tecnologia al servizio di pochi, con ripercussioni dannose per l'intera società.

L'umanesimo digitale rappresenta, dunque, la base per un uso responsabile della tecnologia, garantendo che l'AI e gli altri strumenti siano sempre orientati a potenziare l'uomo, a rispettarne la dignità e a promuovere il benessere collettivo.

(10) J. Nida-Rümelin, N. Weidenfeld, *Umanesimo digitale. Un'etica per l'epoca dell'Intelligenza Artificiale*.

8. *AI come diritto universale*

Nell'era digitale, il diritto di accesso all'intelligenza artificiale rappresenta una naturale evoluzione dell'umanesimo digitale. Così come l'accesso a Internet è stato progressivamente riconosciuto come un diritto fondamentale per partecipare pienamente alla vita sociale, economica e culturale, anche l'intelligenza artificiale deve essere resa disponibile a tutti. L'AI non è solo una tecnologia avanzata, ma un catalizzatore di conoscenza, opportunità e potere decisionale. Il vero progresso si misura non solo in base alle capacità della tecnologia, ma anche in funzione di chi può accedervi e trarne beneficio.

L'umanesimo digitale riconosce che la conoscenza è il fondamento della libertà e dell'emancipazione, e l'AI può svolgere un ruolo cruciale nell'ampliare l'accesso a tali risorse. Dall'istruzione alla sanità, dall'innovazione alle decisioni quotidiane, l'AI offre l'opportunità di trasformare profondamente il nostro modo di apprendere, lavorare e interagire. Tuttavia, affinché rappresenti una vera rivoluzione positiva, deve essere distribuita equamente e accessibile a tutti, indipendentemente da classe sociale, collocazione geografica o disponibilità economica.

Il pericolo è che l'intelligenza artificiale diventi un privilegio, controllato da pochi attori globali o da nazioni più avanzate, aumentando le disuguaglianze già esistenti. Se solo una minoranza potrà sfruttare le risorse e le potenzialità dell'AI, il divario tra chi è incluso e chi resta escluso dal mondo digitale si amplierà, generando nuove forme di esclusione sociale ed economica. Questo rappresenterebbe un fallimento sia per la tecnologia che per la società.

Per evitare tale scenario, è fondamentale riconoscere l'AI come un diritto universale, garantendo che ogni individuo possa beneficiare delle sue opportunità. Non è solo una questione tecnica o economica, ma anche un imperativo etico e sociale. L'intelligenza artificiale deve essere utilizzata per abbattere barriere, non per crearne di nuove. Deve essere uno strumento per democratizzare l'accesso all'informazione, all'istruzione e alle opportunità, colmando i divari sociali e geografici che ostacolano il progresso di intere comunità.

L'accesso equo alla conoscenza è la base su cui si costruisce una società giusta e inclusiva. L'intelligenza artificiale ha il potenziale per eliminare barriere storiche, fornendo strumenti e risorse a chi finora ne è stato escluso. Perché ciò si realizzi, è necessaria una strategia globale che renda l'AI una risorsa condivisa, solo così potrà diventare un motore di crescita e giustizia sociale, concretizzando la visione di un futuro in cui ogni persona, indipendentemente dal contesto in cui vive, possa partecipare e prosperare nel mondo digitale.

* * *

INTELLIGENZA ARTIFICIALE

Una Sintesi Aggiornata

di Salvatore Gaglio (*)

Abstract: L'Intelligenza Artificiale (IA) è in grado di potenziare enormemente le capacità produttive del nostro Paese e può rendere disponibili nuovi servizi. Essa promette, infatti, lo sviluppo di nuovi prodotti che potranno creare nuovi mercati e migliorare la qualità della produzione industriale. Questa tecnologia presenta evidentemente anche dei rischi riguardanti la sicurezza e l'affidabilità dei prodotti di IA, come pure la potenziale disgregazione del mercato del lavoro. Per comprendere appieno l'IA dobbiamo andare alle sue radici. I suoi fondamenti teorici si basano sul concetto formale di *algoritmo* e sono riconducibili al lavoro di Alan M. Turing del 1936, in cui vengono definiti i limiti della calcolabilità e dimostrata l'esistenza di problemi non risolvibili con un calcolo. Una ulteriore intuizione di Turing è quella di una *macchina universale*, ovvero di una macchina in grado di simulare una qualsiasi altra macchina. Su questo concetto si basa, in particolare, il moderno computer *general purpose* programmabile. Questa capacità di simulazione è anche alla base del famoso *Test di Turing*, formulato nel 1950 e spesso citato quando si vuole definire l'IA. L'Italia ha una tradizione di tutto rispetto in questo campo. Già negli anni Sessanta, agli albori dell'IA, in Italia operavano degli istituti di ricerca prestigiosi. E negli anni successivi si è sviluppata una nutrita comunità scientifica di grande rilievo che ha portato avanti progetti di ricerca significativi. Un anno di svolta per l'IA è il 2022, quando OpenAI lancia ChatGPT, un sistema di dialogo in linguaggio naturale, che sfrutta la tecnologia degli LLM (*Large Language Model*) basata sui cosiddetti *Transformer*, ovvero delle reti neurali profonde capaci di catturare le relazioni fra parole di un discorso. Ci si è resi conto che il futuro descritto nei romanzi di fantascienza, come quelli dello scrittore Isaac Asimov, è già oggi. Ma davvero le macchine saranno in grado di sostituirci o di dominarci? Tornando alla realtà, bisogna comprendere quali siano i loro limiti. Un importante limite teorico è noto come problema del riferimento (*symbol grounding*). È il fatto che nessun sistema formale che elabora simboli può determinare univocamente il suo significato. Tuttavia, le macchine possono elaborare significati in modo complesso, anche imprevedibile, ed esibire comportamenti altamente intelligenti, ingannandoci (*Test di Turing*). In conclusione, possono riprodurre tutti gli aspetti calcolabili del pensiero umano (*Macchina Universale di Turing*). L'IA presenta enormi potenzialità tecnologiche, e sicuramente è già in atto una vera nuova rivoluzione industriale che sta cambiando in modo significativo il nostro modo di vivere. Compito della politica è governare rischi e opportunità che essa offre, non lasciandole alle solite logiche liberali di mercato. La politica a tutti i costi deve fare in modo che i vantaggi siano per tutti e i rischi tenuti sotto controllo. Inoltre, la storia ci insegna che la tecnologia si combatte con la tecnologia. Per cui per il nostro Paese, e per l'Europa, l'avvento dell'era dell'IA può rivelarsi invece una grande opportunità di sviluppo, in particolare per il nostro Mezzogiorno.

Artificial Intelligence (AI) has the potential to greatly enhance the productive capacities of our Nation and can make new services available. It promises, in fact, the development of new products that can create new markets and improve the quality of industrial production. This technology also evidently presents risks concerning the safety and reliability of AI products, as well as the potential disruption of the labor market. To fully understand AI, we must go back to its roots. Its theoretical foundations are based on the formal concept of an algorithm and can be traced back to the work of Alan M. Turing in 1936, where the limits of computability were defined and the existence of problems unsolvable by computation was demonstrated. Another insight of Turing is that of a universal machine, that is, a machine capable of simulating any other machine. This concept is particularly the basis of the modern general-purpose programmable computer. This simulation capability is also the basis of the famous Turing Test, formulated in 1950 and often cited when defining AI. Italy has a respectable tradition in this field. Already in the 1960s, at the dawn of AI, prestigious research institutes were operating in Italy. In the following years, a significant scientific community developed, carrying out important research projects. A turning point for AI is 2022, when OpenAI launched ChatGPT, a natural language dialogue system that exploits LLM (Large Language Model) technology based on so-called Transformers, which are deep neural networks capable of capturing the relationships between words in a discourse. It became clear that the future described in science fiction novels, such as those by writer Isaac Asimov, is already today. But will machines really be able to replace or dominate us? Returning to reality, we must understand their limits. An important theoretical limit is known as the symbol grounding problem. It is the fact that no formal system that processes symbols can unambiguously determine their meaning. However, machines can process meanings in a complex, even unpredictable way, and exhibit highly intelligent behaviors, deceiving us (Turing Test). In

(*) Salvatore Gaglio è Professore ordinario di Intelligenza Artificiale presso il Dipartimento di Ingegneria dell'Università degli Studi di Palermo.

conclusion, they can reproduce all the computable aspects of human thought (Turing Universal Machine). AI presents enormous technological potential, and surely a new industrial revolution is already underway, significantly changing our way of life. The task of politics is to govern the risks and opportunities it offers, not leaving them to the usual liberal market logic. Politics must ensure that the benefits are for everyone and the risks are kept under control. Moreover, history teaches us that technology is fought with technology. Therefore, for our country and for Europe, the advent of the AI era can instead prove to be a great development opportunity, particularly for our Mezzogiorno.

Sommario: 1. *L'avvento dell'Intelligenza Artificiale.* – 2. *Le radici teoriche e la storia dell'IA.* – 3. *L'IA italiana.* – 4. *I recenti progressi.* – 5. *Dalla fantascienza alla realtà.* – 6. *Rischi, opportunità, ma prospettive anche per il nostro Mezzogiorno.* – 7. *Ringraziamenti.*

1. *L'avvento dell'Intelligenza Artificiale*

Nel Disegno di Legge sull'intelligenza artificiale predisposto dal Governo Italiano nell'aprile 2024 viene data la seguente definizione di sistema di intelligenza artificiale (termine che da ora in poi indicheremo con la sigla IA):

“Un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali”.

Questa tecnologia è in grado di potenziare enormemente le capacità produttive italiane e di rendere disponibili nuovi servizi. Essa promette, infatti, lo sviluppo di nuovi prodotti che potranno creare nuovi mercati e migliorare la qualità della produzione industriale. I sistemi *software* dotati di tecnologie IA., in grado di prendere o raccomandare decisioni velocemente, possono rendere più efficienti la logistica e le catene di distribuzione, che possono diventare, a loro volta, più robuste adattandosi a eventi non previsti. Nella finanza, le tecnologie IA vengono già utilizzate per formulare per tempo scenari previsivi di rischi diversi e per aumentare la sicurezza nei confronti di frodi e comportamenti malevoli e anomali.

La finanza è proprio uno dei settori principali di impiego delle tecniche di IA. Molte piccole aziende particolarmente dinamiche operano e offrono servizi in vari ambiti della finanza e dell'economia in cui ormai viene impiegata pesantemente l'IA (1), come l'analisi del credito, la gestione del rischio, la *cybersecurity*, il rilevamento delle frodi, la consulenza finanziaria, l'*advertising*, il *trading* e la gestione dei portafogli. Non è un caso quindi che sia aumentato nel mondo anche il venture capital diretto a *startup* in IA. Uno studio della Banca d'Italia del 2022 analizza alcune esperienze del sistema finanziario italiano sull'uso dell'IA nel *credit scoring* (2).

La tecnologia in questione presenta evidentemente anche dei rischi riguardanti la sicurezza e l'affidabilità dei prodotti di IA, come pure la potenziale disgregazione del mercato del lavoro. In un contesto di innovazione digitale orientata solo a incrementare i profitti e a ridurre i costi, essa porta inesorabilmente alla perdita di posti lavoro, che, almeno nel breve periodo, viene compensata solo in piccola misura dalla creazione di nuove opportunità per figure specializzate necessarie allo sviluppo e alla manutenzione di tali sistemi.

Già nelle fabbriche e nella gestione di magazzini, catene robotizzate operano con la supervisione di pochi esseri umani, peraltro sempre elemento ineludibile e decisivo. Per esempio, i leader mondiali dell'e-commerce utilizzano migliaia di robot che si muovono ordinatamente all'interno di grandi spazi opportunamente strutturati per evadere sino a centinaia di migliaia di ordini ogni giorno riguardanti decine di milioni di prodotti. Ma altre figure professionali che fanno parte dei cosiddetti colletti bianchi rischiano di essere sostituiti da sistemi di IA. Agenti *software*, definiti in misura fra l'*utopia* e la distopia come intelligenti, sono in grado di preparare un pezzo giornalistico a partire dalle notizie di agenzia in rete, oppure di inquadrare un caso giuridico tenendo conto della legislazione corrente e delle sentenze precedenti. I sistemi a guida autonoma potrebbero dare luogo a nuove tipologie di trasporto pubblico.

Attualmente ci sono sistemi di IA che vengono autorizzati dall'uomo a prendere decisioni molto importanti per noi. È il caso degli agenti *software* che comprano e vendono azioni in borsa per conto di operatori finanziari che li istruiscono o li addestrano per adottare le loro strategie e per agire con estrema velocità. Muovono volumi pari a trilioni di dollari e contribuiscono a determinare le sorti di aziende e, a volte, nazioni. A un algoritmo di vendita sembra che sia imputabile il crollo della sterlina del 7 ottobre 2016, che è arrivata a perdere il 10% del suo valore in una sorta di effetto domino. Secondo molti, l'algoritmo ha interpretato in senso molto negativo alcune notizie di stampa relative alla Brexit.

Si tratta allora di tecnologie pericolose? Come possiamo fidarci? Si tratta sicuramente di tecnologie sofisticate e potenti, che come tutte le tecnologie avanzate, come quelle ambientali, energetiche, e così via, possono essere

(1) <https://www.money.it/6-modi-Intelligenza-artificiale-rivoluziona-finanza>.

(2) E. Bonaccorsi Patti et al., *Intelligenza artificiale nel credit scoring. Questioni di Economia e Finanza*, n. 721/2022.

usate anche per colpire gli esseri umani, come nel caso dei loro impieghi non controllati in ambito militare. La cosiddetta innovazione digitale va opportunamente progettata considerando fra i criteri importanti l'impatto sociale, così come si fa con altri tipi di progetti in ambito urbanistico, produttivo e di servizi. I processi di innovazione vanno governati opportunamente e la politica non può più ignorare l'emergere di nuovi fenomeni sociali, come, per esempio, la riduzione progressiva della necessità del lavoro umano. Abbiamo visto infatti man mano scomparire agenzie di viaggi, sportelli bancari, uffici e negozi. Si deve fare attenzione a che non tutto venga delegato in modo incontrollato a sistemi digitali. Molte tecnologie di IA riguardano sistemi che possono invece operare in collaborazione con gli esseri umani, aumentandone e migliorandone le capacità.

Ma cos'è davvero l'IA? Per comprenderla appieno dobbiamo partire dalle sue radici teoriche e storiche.

2. Le radici teoriche e la storia dell'IA

I fondamenti teorici dell'IA si basano sul concetto formale di *algoritmo* (3) e sono riconducibili al lavoro di Alan M. Turing del 1936. In esso Turing intende dare una risposta al cosiddetto problema della decisione (*Entscheidungsproblem*) (4) posto dal matematico David Hilbert nel 1928 (5). Il problema riguarda la richiesta di esibire una procedura formale, ovvero un algoritmo, per stabilire se un enunciato della matematica posto nel linguaggio della logica sia vero o falso. Si dice formale una procedura che è espressa in un linguaggio non ambiguo e che è eseguibile in modo meccanico secondo determinate regole. Una prima risposta negativa al problema posto da Hilbert viene data dal logico Alonso Church, subito seguita da quella di Turing.

Per fornire tale risposta, Turing si pone il problema di come formulare una procedura formale, che indichiamo anche con il termine di *algoritmo effettivo*, per dare una soluzione al problema. A quei tempi, gli algoritmi erano ancora espressi con un linguaggio alquanto ambiguo, onde la necessità di formularli in riferimento a un sistema di calcolo automatico. L'idea di Turing è quella di proporre un esperimento mentale. Tali esperimenti erano di moda allora, soprattutto per trovare delle contraddizioni. Per esempio, li usò molto Albert Einstein nelle sue argomentazioni riguardanti i paradossi della fisica quantistica.

Turing pensa a una macchina astratta che può essere realizzata come un sistema meccanico che imita una persona che esegue un calcolo: un essere umano usa un foglio di carta su cui scrive dei simboli che formano delle espressioni e che vengono manipolate per eseguire una procedura di calcolo. Una macchina di Turing realizza in modo meccanico un calcolo di questo tipo attraverso un dispositivo che scrive e legge simboli su un nastro formato da celle contigue (semplificazione del foglio di carta). Il dispositivo può muoversi di una cella a destra o a sinistra a seconda del simbolo letto e di uno stato interno. Gli stati interni, che sono in numero finito, determinano varie fasi dell'esecuzione del calcolo.

Rappresentando con ulteriori simboli gli stati interni, il calcolo effettuato da una macchina di Turing può essere descritto con una collezione di espressioni simboliche che vengono a costituire la descrizione formale e rigorosa di un algoritmo. In questo modo Alan Turing formalizza il concetto generale di *algoritmo*.

Assieme al logico Alonso Church, Turing formula la famosa tesi che costituisce un punto fermo fondamentale nella teoria della calcolabilità: *“un problema ammette una soluzione algoritmica se e solo se esiste una macchina di Turing che lo risolve”*. Questo significa che per dimostrare se un problema sia calcolabile, in altre parole se sia possibile trovare una soluzione, basta fare riferimento a un dispositivo molto semplice come la macchina di Turing, che è facile da trattare in termini matematici.

Infine, una ulteriore intuizione di Turing è il concetto di *macchina universale*. Dato che l'algoritmo realizzato da una macchina di Turing è espresso esso stesso in forma simbolica, può essere codificato nei simboli del nastro di una macchina in grado di interpretarlo e di eseguirlo. È l'idea del moderno computer programmabile che simula ogni altra macchina. Un esempio oggi è il nostro smartphone che a seconda dell'applicazione che esegue diventa telefono, macchina fotografica, rubrica, riproduttore di musica, etc.

La tesi di Church-Turing e il concetto di macchina universale permettono a Turing di esibire un controesempio per il problema della decisione posto da Hilbert. La conclusione è che esistono enunciati della matematica espressi in forma logica non decidibili, per i quali non è possibile avere un algoritmo che stabilisce se siano veri o falsi. In termini generali, esistono problemi non calcolabili per i quali non esiste un algoritmo che possa risolverli. La calcolabilità e la simulazione da parte di una macchina universale rivestono un ruolo fondamentale ed essenziale per stabilire le potenzialità e i limiti dell'IA.

Nei moderni computer l'informazione viene codificata secondo l'aritmetica binaria. L'elemento base è il bit, entità che può assumere uno fra due stati, che vengono indicati rispettivamente con le cifre 0 e 1. I dati vengono pertanto rappresentati simbolicamente da configurazioni di bit. Una singola lettera dell'alfabeto, come, per

(3) Il termine *algoritmo* deriva dal nome del matematico persiano al-Khwarizmi, vissuto nel IX secolo d.c., che nei suoi lavori sviluppò procedure per la soluzione di equazioni matematiche.

(4) Alan M. Turing, On Computable Numbers, with an Application to the Entscheidungsproblem, in *Proceedings of the London Mathematical Society*, ser. 2, vol. 42, 12 novembre 1936, pp. 230-265.

(5) David Hilbert and Wilhelm Ackermann. *Grundzüge der Theoretischen Logik*. Springer, Berlin, Germany, 1928.

esempio, la lettera 'A', viene rappresentata con una configurazione di 8 bit, come '001000001'; un testo come "Nel mezzo del cammin di nostra vita..." viene di conseguenza rappresentato come una lunga sequenza di bit. Così avviene anche per un audio, per un'immagine e per un video. Anche gli stessi programmi sono configurazioni di bit, che vengono opportunamente e meccanicamente interpretati dalla macchina come operazioni da effettuare sui dati. Pure gli algoritmi di IA sono espressi simbolicamente da configurazioni di bit ed effettuano operazioni meccaniche su dati, i quali benché complessi, sono anch'essi espressi come configurazioni di bit. Un computer moderno, da un punto di vista di principio, è quindi solo una realizzazione più efficiente di una macchina di Turing (6).

L'assunzione che sta alla base della ricerca sull'IA è la cosiddetta *Ipotesi del sistema di simboli fisico* di Newell e Simon: (7)

Un sistema di simboli fisico è una macchina, come un computer digitale, che è in grado di manipolare dati simbolici aggiungere numeri, riordinare liste di simboli (come mettere una lista di nomi in ordine alfabetico), sostituire alcuni simboli con altri, e così via. Un sistema di simboli fisico ha i mezzi necessari e sufficienti per agire in modo intelligente. Un aspetto importante è che non importa ciò di cui è fatto il sistema di simboli fisico! Un'entità intelligente potrebbe essere fatta di proteine, materiali meccanici, transistor o di qualsiasi altra cosa, purché sia in grado di elaborare simboli.

Nel 1943 due scienziati, un neurofisiologo e un matematico, Warner McCullochs e Walter Pitts, propongono un modello matematico molto semplice del neurone naturale (8), la singola cellula del sistema nervoso animale. Un neurone artificiale viene attivato se la somma pesata dei suoi ingressi è superiore a una soglia. I pesi modellano le sinapsi fra i neuroni naturali – modulando gli ingressi – e possono essere positivi (simulando sinapsi eccitatorie) o negativi (simulando sinapsi inibitorie). Questi neuroni interconnessi tra loro possono formare delle reti. In tal caso, l'attivazione del neurone viene trasmessa attraverso altre sinapsi ad altri neuroni. McCullochs e Pitts dimostrano che tali reti sono in grado di effettuare operazioni logiche e di riconoscere stimoli visivi.

Nel 1955 un gruppo di scienziati, composto da John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon, propone l'organizzazione di una conferenza a Dartmouth, nel New Hampshire, dal titolo *Dartmouth Summer Research Project on Artificial Intelligence*. L'evento, svoltosi poi nel 1956, viene considerato la nascita della disciplina nota come *Intelligenza Artificiale*. Scopo originale del programma di ricerca era quello di arrivare a un mondo in cui "le macchine usano il linguaggio, formano astrazioni e concetti, risolvono il genere di problemi per adesso riservati agli esseri umani e migliorano loro stesse".

In realtà, alcuni anni prima, ancora Alan Turing nelle due pubblicazioni del 1948 e del 1950, tradotti in italiano rispettivamente con i titoli *Macchine Intelligenti* e *Macchine Calcolatrici e Intelligenza* (9), tratta della meccanizzazione di processi intelligenti, di auto-organizzazione e di apprendimento, nonché del rapporto tra intelligenza umana e intelligenza delle macchine, riprendendo uno dei temi importanti del lavoro del 1936: il concetto di macchina universale.

La possibilità di una macchina di simulare una qualsiasi altra macchina induce Turing a prendere in considerazione la domanda "possono pensare le macchine?". La sua conclusione è che questa domanda è mal posta, perché fa riferimento a concetti ambigui in quel contesto come quelli di "macchina" e "pensare". Pertanto, essa va sostituita con un'altra domanda analoga formulata in termini di un gioco, chiamato "gioco dell'imitazione". Nella sua formulazione originale, a questo gioco prendono parte un uomo (A), una donna (B) e un intervistatore (C). Scopo di A è quello di ingannare C, mentre B cerca di aiutare C. Scopo di C, che non vede direttamente A e B, è capire chi è l'uomo e chi è la donna. In questo contesto la domanda corretta diventa: "Che cosa accadrà se una macchina prenderà il posto di A nel gioco?". È questo il famoso *Test di Turing* – citato in tanti romanzi e film di fantascienza – a cui si fa spesso riferimento come criterio per stabilire se una macchina possa essere considerata intelligente.

Negli anni successivi vengono portati avanti numerosi progetti sull'IA e sviluppate numerose tecniche. Fra i risultati più significativi sono da annoverare il *General Problem Solver* di Simon, Shaw e Newell del 1957 (10), che dà l'avvio alla ricerca sulle tecniche automatiche di *problem solving* con l'uso di euristiche, e il principio di

(6) In realtà un computer fisico ha una limitazione rispetto alla macchina astratta di Turing. Questa infatti prevede un nastro di memoria illimitato, mentre un computer fisico non potrà avere che una memoria finita.

(7) A. Newell, H. A. Simon, Computer Science as Empirical Inquiry: Symbols and Search, *Communications of the ACM*, 19(3), 113-126, 1976.

(8) W. McCulloch, W. Pitts, A Logical Calculus of Ideas Immanent in Nervous Activity, 1943, *Bulletin of Mathematical Biophysics* 5:115-133.

(9) Vedi A. M. Turing, *Intelligenza Meccanica*, a cura di G. Lolli. Bollati Boringhieri, Torino, 1994.

(10) Newell, A.; Shaw, J.C.; Simon, H.A. (1959). Report on a general problem-solving program. *Proceedings of the International Conference on Information Processing*.

risoluzione di Robinson del 1965 (11), che permette la realizzazione di una procedura completa di dimostrazione automatica di teoremi nella logica detta del I ordine.

Nel 2005, per celebrare i 50 anni dell'IA, viene organizzata una nuova conferenza a Dartmouth, con alcuni dei partecipanti alla conferenza del 1956, nel corso della quale sono analizzati i progressi della disciplina. Nel corso degli anni le ricerche sull'IA hanno avuto un grande sviluppo soprattutto nei settori del riconoscimento vocale, dell'elaborazione automatica del linguaggio naturale, della visione artificiale, della risoluzione di problemi, delle tecniche di ragionamento e di apprendimento, della robotica. I successi si sono alternati ai fallimenti. Nella maggior parte dei casi i sistemi realizzati non hanno superato lo stato prototipale. Sembra che sia ancora lontano un sistema di intelligenza artificiale di tipo generale che riproduca l'intelligenza umana in tutti i suoi vari aspetti. Il progresso ha riguardato principalmente la cosiddetta "*Narrow Artificial Intelligence*", che riguarda sistemi specializzati. Poco progresso c'è stato nella "*General Artificial Intelligence*", ovvero un'intelligenza ad ampio spettro di tipo umano.

3. L'IA italiana

L'Italia ha una tradizione di tutto rispetto nel campo dell'IA. Già negli anni Sessanta, agli albori dell'IA, in Italia operano degli istituti di ricerca prestigiosi. Ad Arco Felice, l'Istituto di Cibernetica del Cnr, fondato da Eduardo Caianiello, collabora con i maggiori scienziati di IA di quel tempo e propone alcuni dei primi modelli di reti neurali. A Camogli, l'Istituto di Cibernetica e Biofisica, fondato da Antonino Borsellino, ha anch'esso importanti collaborazioni internazionali e realizza una macchina analogica funzionante, PAPA, per il riconoscimento automatico di figure. In questo istituto si formano anche scienziati di rilievo come Tommaso Poggio e Vincent Torre, che daranno contributi fondamentali allo studio della percezione e alle reti neurali. A partire da questi primi anni presso varie Università e altri Istituti del Cnr fioriscono numerose attività di ricerca nell'IA. Negli anni Settanta, presso il Politecnico di Milano, Marco Somalvico porta in Italia la sua esperienza maturata presso l'Università di Stanford e fonda il "Progetto di Intelligenza Artificiale". Diversi ricercatori universitari e degli istituti del Cnr dell'area pisana portano avanti ricerche in IA in collaborazione con prestigiosi laboratori internazionali. Negli anni Ottanta, Luigia Carlucci Aiello dà avvio alla scuola romana di IA presso "La Sapienza" e dirige i primi progetti nazionali. All'Università di Genova è attivo il "Nucleo Cnr per la Robotica Antropomorfa" con Vincenzo Tagliasco. Molte delle ricerche sulla robotica iniziate presso questo centro verranno poi sviluppate nei nostri giorni presso l'Istituto Italiano di Tecnologia (IIT) di Genova.

Anche l'industria italiana ha prestigiosi trascorsi nell'IA. Negli anni Settanta, l'Elsag di Genova, sotto la direzione di Luigi Stringa, realizza e fornisce alle Poste Italiane il calcolatore parallelo, EMMA, completo di software per il riconoscimento automatico del codice di avviamento postale scritto a mano sulla corrispondenza, che viene adoperato con successo per l'automazione postale.

Nel 1985 Luigi Stringa va a dirigere l'Irst di Trento, che da allora dà grandi contributi all'IA italiana in termini di pubblicazioni e di progetti con l'industria. Successivamente l'istituto prende il nome di Fondazione Bruno Kessler.

Nel 1987 Milano ospita la più importante conferenza internazionale del settore, l'Ijcai, che porta in Italia i maggiori ricercatori internazionali e le tante aziende che in quel periodo sono attive in IA. Anche questo evento contribuisce notevolmente alla crescita della comunità italiana.

Nel corso degli anni 80 vengono anche attivati diversi progetti nazionali sull'IA per conto del Murst e del Cnr. Inoltre, i progetti finalizzati del Cnr degli anni 80 e 90, quali Informatica, Robotica e Sistemi Informatici e Calcolo Parallelo, danno un grande impulso alla ricerca in IA in Italia, con congrui finanziamenti all'accademia e all'industria e con consistenti ricadute in termini di brevetti, prototipi industriali e pubblicazioni scientifiche. Viene anche avviato e concluso uno studio di fattibilità per un progetto finalizzato "Intelligenza Artificiale".

Su queste basi e su queste esperienze pluridecennali è cresciuta una comunità nazionale di ricercatori di grande rilievo internazionale, che operano nell'IA, e sono nate importanti società scientifiche nazionali, quali Girpr (Gruppo Italiano Ricercatori in *Pattern Recognition*), Gulp (Gruppo di Utenti in *Logic Programming*), Glia (Gruppo di Lavoro sull'Intelligenza Artificiale) dell'Aica, AIxIA (Associazione Italiana per l'Intelligenza Artificiale), Siren (Società Italiana Reti Neuroniche). Attualmente circa 900 ricercatori fanno parte del *Artificial Intelligence and Intelligent Systems Lab* (AIIS Lab) del Cini (Consorzio Interuniversitario Nazionale di Informatica), distribuiti in 53 nodi presso varie università italiane ed enti di ricerca. Nel documento "*AI for future Italy*" elaborato dall'AIIS Lab, si rileva che negli ultimi vent'anni l'Italia è nei primi posti per la produzione di lavori scientifici e nelle citazioni e ha inoltre un ruolo centrale nei progetti europei. La comunità italiana dimostra eccellenza internazionale in aree quali il *machine learning*, l'elaborazione del linguaggio naturale, la visione artificiale, la rappresentazione della conoscenza, il ragionamento automatico, la robotica cognitiva, l'IA affidabile, gli agenti mobili e gli algoritmi ad alte prestazioni.

(11) Robinson J.A. "A Machine-Oriented Logic Based on the Resolution Principle". *J. Assoc. Comput. Mach.* 12, 23-41, 1965.

Nell'ambito del Pnrr è stata istituita la fondazione Fair – *Future Artificial Intelligence Research*, costituita da 4 Enti di Ricerca (Cnr, Fondazione Bruno Kessler, Infn, e Iit), 12 Università (Politecnico di Milano, Politecnico di Torino, Sapienza, Scuola Normale Superiore, Università Campus Biomedico di Roma, Università di Bologna, Università di Pisa, Università di Trento, Università di Bari, Università della Calabria, Università di Catania, Università di Napoli “Federico II”) e 5 aziende (Bracco, Expert.ai, Intesa Sanpaolo, Leonardo, Lutech). La fondazione dovrà realizzare un progetto di ricerca nazionale su vari aspetti dell'IA con un finanziamento di 114 milioni di euro.

Centinaia di aziende nazionali, per lo più Pmi, partecipano oggi a eventi nazionali sull'IA come *AI-Week* e *Ital-IA* con prodotti innovativi basati su tecnologie di IA.

4. I recenti progressi

Tra gli anni Ottanta e Novanta c'è un gran fermento intorno all'IA con i cosiddetti *sistemi esperti* (12). Questi sono sistemi progettati per riprodurre il comportamento di un esperto umano, per esempio un medico che fa una diagnosi e assegna una terapia. Nasce la figura dell'*ingegnere della conoscenza*, che ha il compito di realizzare una cosiddetta base di conoscenza in grado di effettuare ragionamenti automatici su un dominio specializzato, realizzato con l'ausilio di un esperto umano del settore. Tali sistemi però raramente superano lo stato di prototipo, perché sono considerati poco flessibili. In quegli anni si rinnova anche l'interesse nei confronti delle reti neurali, costituite da neuroni artificiali interconnessi come quelli di McCulloch e Pitts introdotti in precedenza. Vengono studiate varie architetture con le relative tecniche di autoapprendimento (13).

Grandi progressi si hanno però da 25 anni a questa parte, quando l'IA incontra i sistemi di calcolo ad alte prestazioni, internet e il web, nonché la disponibilità di enormi quantità di dati, i cosiddetti *big data*, su cui poter applicare in modo efficace tecniche statistiche. Già nel 1996 un programma di intelligenza artificiale, *Deep Blue*, sviluppato su un supercomputer dall'Ibm, batte il campione mondiale di scacchi, Kasparov. Sulla scia di *Deep Blue*, l'Ibm realizza il sistema *Watson*, in grado di rispondere a domande poste in linguaggio naturale su vari argomenti. Nel 2011 *Watson* vince un milione di dollari partecipando al gioco a quiz televisivo *Jeopardy* negli Stati Uniti, conquistando il primo posto nei confronti dei suoi avversari umani. Esso costruisce le sue risposte analizzando molto velocemente centinaia di milioni di pagine web su Internet, dalle quali trae le sue conoscenze. Una sua versione specializzata è in grado di fornire risposte rilevanti a problemi di oncologia clinica, individuando anche dei casi molto rari.

Da alcuni anni prototipi sperimentali di auto senza pilota percorrono strade cittadine ed extraurbane, dove sono in grado di individuare i limiti della carreggiata, di percepire il comportamento degli altri autoveicoli e di trovare il miglior percorso per raggiungere la meta, dialogando possibilmente via Internet con altri apparati e veicoli per evitare ingorghi e traffico. Sistemi di visione artificiale riconoscono volti e luoghi in fotografie e impostano automaticamente le modalità di scatto nelle nostre macchine fotografiche. Gli attuali motori di ricerca usano tecniche di IA per rispondere meglio alle richieste, cercando al tempo stesso di comprendere anche le caratteristiche dell'utente con cui interagiscono. I *chatbot* (14), come ChatGPT di *Open AI*, sono in grado di colloquiare in linguaggio naturale con gli esseri umani e di rispondere alle loro esigenze.

Il settore di ricerca che attualmente attira maggiormente l'interesse dei ricercatori è quello del *machine learning*, ovvero dell'apprendimento nelle macchine. L'uso di tecniche statistiche e la disponibilità di enormi quantità di dati (*big data*) prodotti ogni giorno, soprattutto nella rete, ha consentito di mettere a punto dei modelli matematici e dei programmi in grado di autoapprendere dei comportamenti per eseguire dei compiti specifici. Degli agenti *software* definiti intelligenti sono così capaci di selezionare le azioni appropriate in seguito all'individuazione di particolari eventi o di effettuare del *data mining* (scavare nei dati) per scoprire regolarità e correlazioni nascoste.

Uno dei modelli più usati nell'ambito del *machine learning* è quello che si basa sulle reti neurali cui si è accennato in precedenza. Esso cerca di evocare il modo – peraltro assai poco conosciuto – in cui viene elaborata l'informazione nei circuiti nervosi dei sistemi biologici. A somiglianza di questi, le reti neurali artificiali sono realizzate in termini di circuiti che interconnettono in vario modo e con numerosi collegamenti sinaptici dei neuroni artificiali, che, come abbiamo detto, sono versioni semplificate dei neuroni naturali, pur non tenendo conto delle altre tipologie di cellule nervose essenziali per il funzionamento dei neuroni dell'essere umano.

Sono stati studiati diversi modelli di reti neurali a cui corrispondono diversi comportamenti, spesso ispirati ad alcune e limitate funzionalità di sistemi nervosi biologici. L'apprendimento di tali sistemi si basa sulla correzione dell'errore commesso attraverso una modifica locale delle sinapsi, così come sembra potrebbe avvenire nei sistemi naturali, la cui complessità resta da elucidare da parte delle neuroscienze – e questo non è un problema per il quale si possa intravedere una qualche soluzione a breve termine.

(12) Buchanan, B. (1986). “Expert systems: working systems and the research literature”. *Expert Systems*. 3(1): 32-51

(13) David E. Rumelhart, James L. McClelland, PDP Research Group, *Parallel Distributed Processing*, Volume I & II, MIT Press, 1986.

(14) Un *chatbot* è un robot virtuale in rete in grado di dialogare con esseri umani.

Un particolare paradigma di apprendimento nelle macchine è il *deep learning*. Dal 2013 è enormemente (di 6 volte) aumentato il numero di lavori scientifici e brevetti che abbiano a che fare con esso. Si tratta di modelli ispirati – sempre coi limiti di cui sopra – al funzionamento della corteccia cerebrale, in cui strati di neuroni si susseguono per creare rappresentazioni sempre più astratte degli stimoli sensoriali. Numerosi strati di neuroni artificiali creano le loro rappresentazioni interne nascoste per interpretare immagini, linguaggio naturale o sequenze genetiche. Per esempio, traendo ispirazione dagli studi di neurofisiologia, come quelli effettuati sulla corteccia visiva dei gatti da David Hubel e Torsten Wiesel (15), premi Nobel per la Medicina, a partire da immagini digitali vengono rilevati linee con varie orientazioni, contorni, aree omogenee di colore che si aggregano per formare parti significative, per poi individuare e riconoscere oggetti e scene. Tutto questo con processi di apprendimento automatico. Si trovano facilmente in rete – e anche come *app* per smartphone – *deep networks* già addestrate che possono essere utilizzate per impieghi specifici, come nel caso dell'ausilio alla diagnostica medica per immagini.

Le rappresentazioni interne costruite dalle reti neurali profonde risultano di solito non intelleggibili per noi essere umani. Queste entità agiscono da cosiddette scatole nere (*black box*). Riusciamo ad addestrarle per eseguire compiti specifici, ma non comprendiamo del tutto il loro funzionamento interno. Alcune di queste reti sono anche rese accessibili in Internet e continuamente addestrate dagli utenti man mano che esse vengono usate. Il problema della scatola nera preoccupa non pochi. Hanno ottime prestazioni, ma ci possiamo sempre fidare di loro?

Il 2022 è un altro anno di svolta per l'IA con l'avvento di ChatGPT di OpenAI, un sistema di dialogo in linguaggio naturale, che sfrutta la tecnologia degli Llm (*Large Language Model*) basata sui cosiddetti *Transformer*, ovvero delle reti neurali profonde capaci di catturare le relazioni fra parole di un discorso. ChatGPT si rivela in grado di trattare qualsiasi argomento con un interlocutore umano in diverse lingue. Il sistema è addestrato con la tecnica dell'apprendimento profondo (*deep learning*) tramite l'enorme mole di conoscenza contenuta nel *web*. Il suo funzionamento segue l'approccio cosiddetto *generativo*, che consiste nella generazione di frasi di dialogo opportunamente contestualizzate. Tale approccio è stato poi generalizzato facendo sì che tali sistemi possano generare anche altri tipi di contenuti, come immagini, video, audio e anche codice *software* funzionante.

L'avvento dei computer quantistici permetterà di avere macchine superveloci in grado di risolvere problemi finora inaccessibili anche ai più potenti supercomputer. Google, Nasa e alcune Università statunitensi hanno fondato il *Quantum Artificial Intelligence Lab* in California, con un programma di ricerca volto ad esplorare le nuove potenzialità che questa tecnologia offre. Anche Microsoft sta puntando sul *quantum computing* con il suo nuovo centro di ricerca *Station Q*, mentre Intel sta progettando processori con capacità di elaborazione quantistica tramite l'uso di nuovi tipi di transistor. L'Ibm mette a disposizione degli sviluppatori dei server quantistici tramite l'ambiente di sviluppo *qiskit*.

I computer quantistici possono operare su scala microscopica, permettendo allo stesso tempo di far evolvere diverse linee di calcolo in sovrapposizione sugli stessi bit (detti *qubit*). In questo modo possono dare luogo a un salto esponenziale in termini sia di memoria sia di velocità di calcolo. I problemi da risolvere dal punto di vista tecnologico sono ancora tanti, ma si può già cominciare a sperimentare. Essi sarebbero particolarmente adatti per le tecniche più simboliche di IA, in particolare nell'ambito della risoluzione automatica di problemi, che finora hanno avuto applicazione limitata in quanto i computer tradizionali, anche i più potenti, non sono in grado di gestire la grande richiesta di calcolo che cresce in modo esponenziale con le dimensioni dei problemi da risolvere.

5. Dalla fantascienza alla realtà

Solo negli ultimi anni, e soprattutto con l'avvento dell'IA generativa come ChatGPT, basata sul paradigma del *deep learning* con l'uso di *big data*, ci si è resi conto che il futuro descritto nei romanzi di fantascienza, come quelli dello scrittore Isaac Asimov, è già oggi.

Asimov nella sua saga dedicata ai robot tratta proprio del rapporto tra uomo e robot e dei pericoli dell'IA, per cui formula le sue famose tre leggi della robotica intese a limitare il potere delle macchine. Introdotte per la prima volta nel suo racconto breve intitolato "Runaround", pubblicato nel 1942, esse recitano così:

1. Un robot non può recare danno a un essere umano, né può permettere che, a causa della propria inazione, un essere umano riceva danno.

2. Un robot deve obbedire agli ordini impartitigli dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge.

3. Un robot deve proteggere la propria esistenza, purché ciò non contravvenga alla Prima o alla Seconda Legge.

Nel racconto "Reason" del 1941, e poi in altri racconti e romanzi successivi, Asimov tratta anche di un calcolatore universale, chiamato *Multivac*. *Multivac* è un supercomputer che rappresenta una forma avanzata di IA. Esso viene utilizzato in diverse storie per esplorare i temi legati alla gestione dell'informazione, alla previsione del futuro e al controllo delle decisioni umane. Viene immaginato come un'entità onnipotente capace di elaborare enormi quantità di dati e di risolvere complessi problemi scientifici e sociali. Sembra che si parli di ciò che si

(15) HUBEL DH, WIESEL TN. Receptive fields of single neurones in the cat's striate cortex. *J Physiol*. 1959 Oct;148(3):574-91

prospetta oggi con sistemi già in uso come ChatGpt, o Bard, o Gemini, o Claude, etc. Il *cervello positronico* dei robot di Asimov fa poi pensare al funzionamento come scatole nere delle reti neurali profonde.

Ma davvero le macchine saranno in grado di sostituirci o di dominarci? Tornando alla realtà, vediamo di comprendere anche quali siano i loro limiti. Come si è detto, l'intelligenza artificiale si realizza con algoritmi che fanno sì che le macchine elaborino informazione simbolica. Gli algoritmi stessi sono formulati con espressioni simboliche. Dell'informazione si riesce a dare soltanto una misura quantitativa. La quantità di informazione contenuta in un dato o in un messaggio è stata definita dal cibernetico Claude Elwood Shannon (1916-2001) in termini del logaritmo della sua probabilità di presentarsi e corrisponde al numero di *bit* necessario per codificarlo (16). Non è invece possibile attribuire ai simboli un significato preciso e univoco.

È il cosiddetto problema del riferimento (*symbol grounding*). Nessun sistema formale può determinare univocamente il suo significato. Una funzione di interpretazione dei simboli non può essere definita esclusivamente in termini di altri simboli, perché ciò darebbe luogo a una regressione infinita. Un esempio ci è dato dalla difficoltà di interpretazione delle antiche scritture. Abbiamo bisogno di una chiave di lettura che ci dica come interpretare i simboli di una lingua. Per i geroglifici egiziani l'ufficiale napoleonico Champollion (1790-1832) la trovò nella stele di Rosetta, grazie alla traduzione di un testo egiziano in testi in lingue conosciute.

Il filosofo John Searle crede che ciò di cui siamo fatti sia fondamentale per la nostra intelligenza. Il pensiero può esistere soltanto in macchine molto speciali quelle viventi fatte di proteine. Il suo aneddoto della stanza cinese (17) tende a dimostrare che in linea di principio sia possibile simulare in modo meccanico, così come su un calcolatore, un comportamento intelligente, senza che però ci sia alcuna comprensione o attività di pensiero. L'aneddoto riguarda la realizzazione di un sistema in grado di comunicare in una lingua straniera, come, per esempio, il cinese, realizzato con un essere umano che non conosce tale lingua, ma che all'interno di una stanza riceve dall'esterno una sequenza di schede con dei simboli in cinese e, seguendo delle regole fisse che può leggere in un manuale, costruisce a partire dalle schede in ingresso una risposta, sempre in cinese, recuperando da opportuni schedari all'interno della stanza una nuova sequenza di schede che costituiscono la risposta. Se il manuale che l'essere umano consulta è abbastanza ricco di regole, ecco che abbiamo ottenuto un sistema in grado di colloquiare, anche se con delle schede, in cinese. Ma Searle si chiede: dove sta l'attività di pensiero nell'atto di colloquiare? Il manuale e le schede sono oggetti inerti. È vero che c'è l'essere umano, ma questi non conosce il cinese; segue solo delle regole sui simboli da recuperare a partire da quelli ricevuti e non sa di cosa si stia parlando.

Una caratteristica di noi esseri umani è quella di essere esseri senzienti. Riusciamo a dare un significato ai simboli perché abbiamo quella che si dice un'esperienza cosciente soggettiva. Le sensazioni, quelle che i filosofi chiamano *qualia*, sono quelle che noi proviamo con le nostre percezioni, che non sono solo di natura simbolica. Tuttavia, con i metodi simbolici si riesce a elaborare anche in maniera molto sofisticata i significati attraverso metodi detti compositivi. A partire da simboli primitivi a cui noi, esseri umani, attribuiamo un significato preciso, si possono elaborare delle espressioni simboliche che rappresentano concetti sempre più complessi in modo coerente ai significati che diamo noi.

La nostra percezione di un colore, come la qualità di rosso che vediamo, è diversa da ciò che succede in un sistema di IA. È chiaro che una macchina è in grado di discriminare i colori attraverso sensori che misurano la lunghezza d'onda della luce riflessa dagli oggetti, ed è anche in grado di farlo meglio di noi umani. Tale discriminazione però si basa su una traduzione della misura in simboli che vengono tra loro confrontati. La sensazione di rosso, verde, blu che noi esseri umani abbiamo nella nostra esperienza cosciente è tutta un'altra cosa. Così come le sensazioni di dolore, piacere, gioia, tristezza, melodia, etc. sono parte della nostra esperienza cosciente. La macchina può simularle, ma non provarle. Così come un algoritmo può simulare un ponte in un sistema di Cad di ingegneria, ma il ponte vero è un'altra cosa.

Caratteristica della mente umana è l'unitarietà e la soggettività dell'esperienza cosciente: l'esperienza è un tutto contemporaneamente, riguarda il singolo individuo ed è inaccessibile ad altri. Proprio in quest'esperienza i simboli trovano il loro significato. Il funzionamento globale di una macchina è invece completamente spiegabile dall'esterno in termini riduzionistici, tramite parti tra loro interagenti che elaborano espressioni simboliche. Descrive bene questo aspetto Federico Faggin, inventore del microprocessore: "*La macchina tipica è un sistema riduzionista interamente costruita da parti separabili che possono essere smontate e poi rimontate per ripristinarne in pieno il funzionamento*" (18). Il suo comportamento è totalmente determinato da configurazioni di bit (simboli) di cui sono fatti i suoi programmi e i suoi dati. Come si è detto, i simboli da soli non catturano il significato. Continua Faggin: "*All'interno di un computer, l'intenzione soggettiva e il significato del programmatore sono stati interamente oggettivati nel programma. Il risultato dell'elaborazione meccanica eseguita dal computer viene infine restituito alla mente umana per essere interpretato soggettivamente.*"

(16) C. E. Shannon e Warren Weaver, *La teoria matematica delle comunicazioni*, Etas Kompass, Milano, 1971.

(17) J. Searle, *Menti cervelli e programmi, un dibattito sull'intelligenza Artificiale*, CLUP-CLUED, 1984.

(18) F. Faggin, *Irriducibile*. Mondadori, Milano 2022.

In sintesi, possiamo dire che le macchine, anche se sono spesso molto più efficienti degli esseri umani, sono limitate rispetto alla mente umana nei seguenti aspetti:

- le macchine elaborano solo simboli. Il loro comportamento è determinato in ogni istante dalla configurazione di bit in memoria (stato computazionale);
- le macchine non hanno sensazioni (di colore, di suoni, di dolore, di tristezza, di gioia, di brama, di melodia, ecc.), non hanno esperienza cosciente, la loro coscienza – ove mai di desse una definizione di coscienza – può essere solo “artificiale”;
- le macchine non possiedono il significato dei simboli che elaborano, sono solo automi, e senza significati non possono avere volontà propria;
- l'esecuzione di un algoritmo (sequenza di simboli) determina solo l'evoluzione temporale degli stati di una macchina, quindi, il software non può mai dare luogo a una mente (altro “oggetto” difficile da definire) di un essere senziente.

Tuttavia, le macchine possono elaborare significati in modo complesso, anche imprevedibile, ed esibire comportamenti altamente intelligenti, ingannandoci (Test di Turing). In conclusione, possono riprodurre tutti gli aspetti calcolabili del pensiero umano (Macchina Universale di Turing).

La tecnologia delle reti neurali profonde inoltre è in grado di fornire prestazioni sorprendenti, anche se si realizzano sistemi che non hanno alcuna comprensione. La presenza di tanti strati di neuroni (da cui il nome di reti profonde) permette loro di costruire rappresentazioni interne delle informazioni che elaborano man mano più astratte e in genere per noi incomprensibili (da cui la qualifica di *black box* che viene loro data). Una rete come ChatGPT, con questa tecnica, costruisce le frasi aggiungendo, una dopo l'altra, parole sulla base della loro probabilità condizionata al contesto del discorso e ai dati con cui è stata addestrata. Quest'operazione è puramente sintattica e risultato di un calcolo, non c'è dietro alcuna comprensione, ma è sorprendentemente coerente con i significati che noi attribuiamo alle frasi.

Il modello di rete neurale che viene realizzato all'interno della macchina risulta essere alla fine una fitta rete di operazioni aritmetiche, per lo più addizioni e moltiplicazioni, attraverso le quali fluiscono i dati a partire dagli ingressi. Tanto è vero che i processori utilizzati per il calcolo sono delle Gpu (*Graphic Processor Unit*), che sono state sviluppate in precedenza per applicazioni grafiche, che per loro natura richiedono grandi quantità di operazioni di moltiplicazione e addizione. La maggiore produttrice di Gpu, *Nvidia*, è diventata infatti leader mondiale per la produzione di sistemi di calcolo per l'IA.

Una rete neurale di questo tipo è definita da un numero enorme di parametri (dell'ordine dei 1000.000.000.000), che sono le sue sinapsi, per cui per noi è una vera e propria *black-box*. Il suo comportamento non-spiegabile, determinato complessivamente dai valori di tutti questi parametri, è anche il suo punto di forza. L'algoritmo di apprendimento esplora uno spazio di soluzioni di dimensioni pari al numero dei suoi parametri (1000.000.000.000), quindi per noi umani assolutamente inimmaginabile.

È stato dimostrato che una rete neurale con almeno tre strati e un numero sufficiente di neuroni per strato è anche un approssimatore universale di funzioni (19). In questo caso la funzione è la relazione tra ingresso e uscita che si vuole fare apprendere alla rete con gli esempi forniti con l'addestramento. Ciò vuol dire, quindi, che con un numero adeguato di esempi è possibile fare apprendere alla rete una qualsiasi funzione. Come si è detto, essa esplora uno spazio di enormi dimensioni ed è pertanto in grado di generare soluzioni a cui noi non potremmo neppure pensare.

6. Rischi, opportunità, ma prospettive anche per il nostro Mezzogiorno

Abbiamo visto come l'IA presenti enormi potenzialità tecnologiche, e sicuramente è già in atto una vera nuova rivoluzione industriale che sta cambiando in modo significativo il nostro modo di vivere. Compito della politica è governare rischi e opportunità che essa offre, non lasciandole alle solite logiche liberali di mercato. La politica a tutti i costi deve fare in modo che i vantaggi siano per tutti e i rischi tenuti sotto controllo.

Il Parlamento Europeo ha da poco approvato l'*AI Act* che intende dare delle direttive ai governi nazionali perché questi legiferino al fine di tenere sotto controllo l'IA. In esso vengono in particolare classificati i rischi di tale tecnologia. In primo luogo, ci sono i rischi inaccettabili:

“L'elenco delle pratiche vietate di cui al titolo II comprende tutti i sistemi di IA il cui uso è considerato inaccettabile in quanto contrario ai valori dell'Unione, ad esempio perché viola i diritti fondamentali. I divieti riguardano pratiche che presentano un elevato potenziale in termini di manipolazione delle persone attraverso tecniche subliminali, senza che tali persone ne siano consapevoli, oppure di sfruttamento delle vulnerabilità di specifici gruppi vulnerabili, quali i minori o le persone con disabilità, al fine di distorcerne materialmente il comportamento in maniera tale da provocare loro o a un'altra persona un danno psicologico o fisico La proposta vieta altresì l'attribuzione di un punteggio sociale basato sull'IA per finalità generali da parte di

(19) Hornik, Kurt; Stinchcombe, Maxwell; White, Halbert (January 1989). “Multilayer feedforward networks are universal approximators”. *Neural Networks*. 2 (5): 359-366.

autorità pubbliche. È infine vietato anche il ricorso a sistemi di identificazione biometrica remota “in tempo reale” in spazi accessibili al pubblico a fini di attività di contrasto, fatta salva l'applicazione di talune eccezioni limitate”.

Vengono poi individuati i sistemi ad alto rischio. Fanno parte di questa categoria le applicazioni che riguardano le infrastrutture critiche, l'istruzione e la formazione professionale, l'occupazione, i servizi pubblici e privati (assistenza sanitaria, banche), i sistemi di contrasto, migrazione, giustizia e processi democratici (per esempio, i sistemi usati per influenzare le elezioni). Per questi sistemi c'è l'obbligo di valutare e ridurre i rischi. Essi dovranno avere requisiti di trasparenza nel rispetto delle norme UE sul diritto d'autore riguardo ai dati usati per l'addestramento dei vari modelli. Inoltre, i *deepfake*, ovvero testi, immagini, audio e video manipolati dovranno essere chiaramente etichettati.

D'altra parte, l'IA è una tecnologia che offre grandi vantaggi, alcuni dei quali vengono suggeriti dalla stessa ChatGPT e riguardano i seguenti campi:

1. *Automazione dei compiti ripetitivi: L'IA può automatizzare attività ripetitive e di routine, consentendo alle persone di concentrarsi su compiti più creativi, complessi e strategici. Ciò porta a un aumento dell'efficienza e della produttività.*

2. *Analisi dei dati avanzata: L'IA è in grado di analizzare grandi quantità di dati in modo più rapido ed efficiente rispetto agli esseri umani. Ciò consente di ottenere insight approfonditi e di prendere decisioni informate basate sui dati.*

3. *Apprendimento continuo: I modelli di IA possono essere addestrati su grandi dataset e migliorare continuamente con l'esperienza. Questo permette loro di adattarsi a nuovi scenari e cambiamenti nell'ambiente senza richiedere una riconfigurazione completa.*

4. *Riduzione degli errori umani: L'IA può contribuire a ridurre gli errori umani causati dalla stanchezza, dalla distrazione o da fattori emotivi. Questo è particolarmente importante in settori come la medicina e l'industria, dove errori possono avere conseguenze gravi.*

5. *Personalizzazione e raccomandazioni: L'IA può analizzare i comportamenti e le preferenze degli utenti per offrire raccomandazioni personalizzate. Questo si traduce in un'esperienza utente migliorata su piattaforme come i motori di ricerca, i servizi di streaming e gli shop online.*

6. *Automazione industriale: Nell'ambito della produzione e della logistica, l'IA può ottimizzare i processi di produzione, gestire la catena di approvvigionamento e migliorare la manutenzione predittiva delle attrezzature.*

7. *Assistenza sanitaria avanzata: L'IA può aiutare i professionisti sanitari nella diagnosi precoce di malattie, nella personalizzazione dei piani di trattamento e nell'analisi di grandi dataset per individuare modelli medici.*

8. *Veicoli autonomi: L'IA è fondamentale per lo sviluppo dei veicoli autonomi, migliorando la sicurezza stradale e offrendo nuove opportunità nel settore dei trasporti.*

9. *Linguaggio naturale e interazione uomo-macchina: I sistemi di IA che comprendono e generano il linguaggio naturale consentono una migliore interazione tra gli utenti e le macchine, facilitando l'accesso alle informazioni e migliorando l'usabilità.*

10. *Risparmio di tempo e risorse: L'automazione fornita dall'IA può portare a un notevole risparmio di tempo e risorse, contribuendo a ridurre i costi operativi e aumentando l'efficienza in vari settori.*

Recentemente il Governo Italiano ha predisposto un Disegno di Legge in materia di IA che intende basarsi su una visione detta antropocentrica.

Nel Ddl vengono evidenziati alcuni requisiti che devono avere i sistemi di IA, per esempio esso recita:

- *Al fine di garantire il rispetto dei diritti di cui al presente articolo deve essere assicurata, quale preconditione essenziale, la cybersicurezza lungo tutto il ciclo di vita dei sistemi e dei modelli di intelligenza artificiale....*

- *Le informazioni e le comunicazioni relative al trattamento dei dati connesse all'utilizzo di sistemi di intelligenza artificiale avvengono con linguaggio chiaro, semplice ed esaustivo, in modo da garantire all'utente la piena conoscibilità e la facoltà di opporsi ai trattamenti non corretti dei propri dati personali.*

Vengono indicati anche alcuni ambiti di intervento, come i seguenti:

- *Lo Stato e le altre autorità pubbliche ... promuovono l'utilizzo dell'intelligenza artificiale come strumento per migliorare l'interazione uomo-macchina nei settori produttivi e migliorare la produttività in tutte le catene del valore e le funzioni organizzative, nonché quale strumento utile all'avvio di nuove attività economiche, al fine di accrescere la competitività del sistema economico nazionale e la sovranità tecnologica della Nazione nel quadro della strategia europea;*

- *L'utilizzo di sistemi di intelligenza artificiale contribuisce al miglioramento del sistema sanitario e alla prevenzione e cura delle malattie, nel rispetto dei diritti, delle libertà e degli interessi dell'interessato, anche in materia di protezione dei dati personali.*

L'aspetto antropocentrico è centrale anche nell'intervento "*Intelligenza Artificiale e Pace*" di Papa Francesco alla LVII Giornata Mondiale della Pace (20) l'1 gennaio 2024. Il Santo Padre parla del progresso della scienza e della tecnologia come via verso la pace, trattando del futuro dell'intelligenza artificiale tra promesse e rischi, del senso del limite nel paradigma tecnocratico, dei temi scottanti per l'etica, delle sfide per l'educazione e delle sfide per lo sviluppo del diritto internazionale, concludendo con atteggiamento positivo e con un auspicio: "Trasformeremo le spade in vomeri?". Tali temi vengono ripresi da papa Francesco al G7 nel giugno 2024, intervenendo ancora sul pericolo dell'uso dell'IA in campo militare e sulla necessità di far sì che le decisioni vengano sempre prese dagli esseri umani.

Tuttavia, un atteggiamento esclusivamente difensivo nei confronti dell'intelligenza artificiale, come sembra profilarsi in gran parte delle iniziative a livello nazionale ed europeo, non potrà essere affatto sufficiente ed efficace. Se non diventiamo competitivi anche noi italiani ed europei, ci toccherà subirla, perché altre nazioni, *in primis* gli USA, la Cina, la Russia, e ora anche l'India, continueranno a svilupparla ed a usarla a loro esclusivo vantaggio. Inoltre, questa tecnologia è nelle mani di potenti multinazionali come Google, Microsoft, Amazon, etc.

La storia ci insegna che la tecnologia si combatte con la tecnologia. Anche per il nostro Paese e per l'Europa l'avvento dell'era dell'IA può essere invece una grande opportunità, in particolare per il nostro Mezzogiorno.

Anche le Università del Sud Italia, e ormai anche le scuole secondarie, formano giovani con competenze diffuse in IA (vedi percorsi di cosiddetta curvatura), secondo diversi approcci che vanno dalla pura tecnologia, ai suoi usi e alle sue ricadute sociali ed economiche. Se non si fa nulla, questi giovani, per sfruttare le loro competenze, saranno sempre di più costretti ad emigrare, con il risultato che noi continueremo a spendere le nostre risorse umane e finanziarie per far avanzare e arricchire altre realtà.

Da dati di Confindustria e Anitec-Assinform del 2023(21) risulta che nel Mezzogiorno la percentuale di aziende che utilizzano sistemi di IA è in linea con la media europea (7,6%), a differenza della media nazionale che è appena il 6,2%.

Perché, quindi, non rendere competitivo il nostro Mezzogiorno con un intervento pubblico-privato che costituisca poli di eccellenza nella ricerca applicata in questo settore – sul modello di quelli esistenti in Germania, in Inghilterra, in Francia e nei Paesi Bassi – da cui far nascere tanti *spin-off*? Le applicazioni potrebbero stimolare altri settori come la sanità, il turismo, i trasporti, l'agricoltura e nuove forme di artigianato. Il nostro Mezzogiorno potrebbe così cavalcare l'onda dell'IA, diventando più ricco senza veri danni ambientali, con la creazione di posti di lavoro, grazie anche a un indotto, e con la diminuzione del devastante fenomeno dell'emigrazione dei nostri giovani.

La robotica potrebbe aiutare il settore agricolo, che in questo momento è in crisi, con sistemi automatizzati di monitoraggio, gestione, pianificazione e raccolta, inseriti nelle coltivazioni. Il turismo potrebbe godere di nuovi servizi avanzati di assistenza ai turisti, che spesso hanno difficoltà a muoversi nel territorio. Potremmo migliorare le prestazioni dei servizi pubblici, che sappiamo essere molto carenti, a cominciare dalla sanità, dai trasporti, dalla gestione dei rifiuti e dalla pubblica amministrazione in generale. Attività di monitoraggio aiuterebbero a preservare e a migliorare l'ambiente. A questi si potrebbero affiancare dei robot in grado di pulire spiagge e mari, o sistemi di sorveglianza per prevenire il dramma degli incendi. Un ulteriore settore di intervento molto importante sarebbe quello della logistica, che l'IA potrebbe rendere efficiente. Nulla poi ci vieta di sviluppare prodotti innovativi, legati anche all'artigianato, da esportare fuori dal territorio.

È chiaro che bisogna vincere le resistenze burocratiche e corporative con politiche più lungimiranti. Le risorse da finanziare sarebbero principalmente quelle umane e quelle infrastrutturali. Dovrebbe però essere garantito un canale di finanziamento costante, altrimenti si ripeterebbero gli errori del passato. Le tecnologie di per sé non avrebbero un elevato impatto ambientale, lo sviluppo di sistemi di IA richiederebbe per lo più l'impiego di cervelli da far lavorare assieme, che però oggi ci prendiamo il lusso di esportare altrove. I *server* ad alte prestazioni che sarebbero richiesti per il calcolo dovrebbero essere gestiti centralmente a livello nazionale come servizi in *cloud*, anche per gli elevati consumi energetici, assieme a depositi nazionali condivisi di *big data* opportunamente certificati.

Bisogna però evitare un ulteriore rischio molto subdolo, quello della banalizzazione dell'IA. Ambienti di sviluppo di sistemi basati sul paradigma del *machine learning* sono ormai quasi alla portata di chiunque conosca un po' di programmazione. Esistono librerie collegate a vari linguaggi e ambienti software con cui è facile costruire un'applicazione di IA, avendo a disposizione dati in sufficienza per l'addestramento. Si ripresenta ancora una volta, e non a caso, un fenomeno che abbiamo visto col diffondersi dei *Personal Computer*. Chiunque ha saputo mettere le mani su un calcolatore si è spacciato per informatico. Il risultato è stato il fallimento di un enorme numero di progetti di informatizzazione e, soprattutto, l'estrema vulnerabilità dei sistemi realizzati.

(20) <https://www.vatican.va/content/francesco/it/messages/peace/documents/20231208-messaggio-57giornatamondiale-pace2024.html>

(21) <https://www.innovationpost.it/attualita/il-digitale-motore-di-sviluppo-per-il-paese-nel-2023-21-a-quota-79-miliardi-nei-prossimi-anni-crescita-media-del-39-annuo/>

Anche la realizzazione di sistemi di IA richiede adeguate professionalità, che devono essere maturate con anni di studio e di esperienza. Non basta assemblare componenti software pronti per l'uso che le grandi multinazionali mettono a disposizione, bisogna saperci guardare dentro e realizzare anche le proprie soluzioni. I professionisti dell'IA devono poi conoscere bene le tecniche informatiche avanzate e, in particolare, quelle relative alla *cyber-security*, per far sì che i loro sistemi non siano facilmente attaccabili da *hacker* che potrebbero volgere a loro vantaggio la loro operatività o addirittura bloccarla.

A ciò si collega un ulteriore motivo di criticità che desta non poche preoccupazioni. Anche nella pratica informatica si sta diffondendo a dismisura l'uso dell'IA generativa per produrre software. Si può chiedere infatti a uno dei *chatbot* attualmente accessibili di scrivere del codice per un'applicazione. Tale uso velocizza enormemente i processi di produzione in un'azienda informatica. Tuttavia, poiché il software prodotto si basa su una vasta varietà di programmi forniti per l'addestramento, è facile che i difetti (*bug*) già presenti si propaghino in maniera esponenziale. Chiaramente il *software* prodotto deve essere controllato e validato da professionisti informatici, ma quando esso raggiunge dimensioni come migliaia o milioni di righe di codice, chi ne garantisce la correttezza? Anche questa è una sfida che richiede professionalità adeguate, per cui esistono davvero molte opportunità di innovazione per impiegare i nostri migliori cervelli nei nostri territori.

Un ultimo aspetto riguarda gli effetti di questa tecnologia sull'ambiente e la sua sostenibilità. Non si può certo dire che essa sia in accordo con le strategie *Green*. In particolare, l'addestramento di reti neurali profonde con *big data* richiede l'impiego di potenti server che consumano enormi quantità di energia, e ciò incoraggia ancora il ricorso alle fonti fossili per generare la corrente elettrica necessaria ad alimentare le macchine. Francesco Beltrame nel suo articolo "La Termodinamica dell'Intelligenza Artificiale e lo Spazio" (22) riporta dei dati impressionanti riguardo a quanto venga a costare dal punto di vista energetico l'addestramento di una IA generativa. Si prospettano scenari in cui il fabbisogno energetico dell'IA possa superare la quantità di energia che l'umanità è in grado di produrre. Ancora un'altra sfida che i nostri progettisti potranno avere l'opportunità di affrontare.

7. Ringraziamenti

Un ringraziamento a Francesco Beltrame per i suoi preziosi commenti e suggerimenti.

* * *

(22) F. Beltrame, La termodinamica dell'Intelligenza Artificiale e lo Spazio. *Agenda Digitale*, in corso di stampa.

DISINFORMAZIONE E INTELLIGENZA ARTIFICIALE: UN COCKTAIL ESPLOSIVO?

di Oreste Pollicino (*)

Abstract: L'intelligenza artificiale ha permeato i diversi aspetti della società moderna, compreso quello della diffusione delle informazioni. Se da un lato l'IA offre un immenso potenziale per migliorare l'accesso e l'elaborazione delle informazioni, dall'altro si è posta come strumento idoneo a generare e diffondere disinformazione, rappresentando una minaccia significativa per i processi democratici e il dibattito pubblico. Questo contributo intende approfondire il ruolo multiforme che svolge l'IA nella lotta alla disinformazione, soffermandosi sul suo coinvolgimento nella creazione, nella diffusione e nel rilevamento dei contenuti falsi e analizzando le possibili prospettive future.

Artificial intelligence has permeated various aspects of modern society, including the realm of information dissemination. While AI offers immense potential to enhance information access and processing, it has also emerged as a tool for generating and spreading disinformation, posing a significant threat to democratic processes and public discourse. This paper delves into the multifaceted role of AI in the fight against disinformation, examining its involvement in content creation, dissemination, and detection and analyzing the future perspectives.

Sommario: 1. *Introduzione: i parametri costituzionali rilevanti.* – 2. *Disinformazione, misinformazione, fake news: profili definitori.* – 3. *Intelligenza artificiale e produzione della disinformazione.* – 4. *Intelligenza artificiale e diffusione della disinformazione.* – 5. *Conclusioni: quale futuro per il governo della disinformazione?*

1. *Introduzione: i parametri costituzionali rilevanti*

Da un punto di vista costituzionale, il tema del possibile contrasto alla diffusione della disinformazione inter-cetta l'essenza del paradosso di Böckenförde, secondo il quale «lo Stato liberale secolarizzato vive di presupposti che non può garantire». È proprio sul modello di tutela prescelto della libertà di espressione – una delle cifre distintive dello stato liberale – che si gioca infatti la partita decisiva in quest'ambito. La scelta di intervenire per depurare la rete dalla disinformazione sembrerebbe evocare infatti, secondo alcuni, l'adozione di una serie di limitazioni che costringono la libertà di parola entro limiti probabilmente più rigorosi rispetto a quelli codificati dalle costituzioni liberali: così, per evitare il propagarsi di odio e falsità online. Dunque, in nome della protezione di diritti costituzionalmente rilevanti, si potrebbe giungere a limitare in maniera rilevante una libertà.

Basterebbero queste primissime osservazioni per rilevare come le ragioni che più recentemente sono comparse sulla scena del dibattito politico rispetto alla necessità di introdurre strumenti normativi per contrastare il propagarsi di *fake news* e discorsi d'odio, e i conseguenti effetti dannosi per l'opinione pubblica, tocchino direttamente i concetti di parole e potere. Ciò rende necessario per chiunque intenda addentrarsi in questo dibattito, oltre la patina di un'analisi meramente superficiale, confrontarsi con lo statuto costituzionale della libertà di espressione. Un paradigma complesso e differente a seconda del sistema giuridico di riferimento, pur nella comune matrice liberale che contraddistingue l'Europa e gli Stati Uniti (1). Volendo semplificare, si potrebbe rappresentare una dicotomia che oppone, rispetto alla circolazione di idee e opinioni, l'opzione di una democrazia militante, impegnata nella strenua difesa di un sistema valoriale, che l'esercizio degli stessi diritti costituzionalmente tutelati rischia di mettere a repentaglio (2), e il modello di una democrazia tollerante, dove l'idea di uno Stato etico o comunque di un maggior controllo sull'esercizio delle libertà da parte degli individui appare svanire. Queste diverse sensibilità oggi sono sottoposte sempre più a confronto, alla luce dell'avvento di un mezzo come Internet che ha permesso di collegare utenti con loro idee, le loro opinioni e le loro manifestazioni di pensiero sparsi in diverse parti del mondo. Sovente le discussioni che coinvolgono attivisti della rete e organizzazioni non governative evocano o l'opzione di un autogoverno della rete (tendenzialmente recessiva, alla luce delle ormai dimostrate scarse capacità dei poteri pubblici di interferire con il funzionamento di Internet), oppure la possibilità di creare carte dei diritti sovranazionali, le quali, tuttavia, sono di difficile riconciliazione con il pluralismo costituzionale che contraddistingue alcuni ambiti, primo su tutti la libertà di espressione. Il confronto con il diverso paradigma

(*) O. Pollicino è professore ordinario di diritto costituzionale, Università Bocconi e rappresentante italiano presso l'Agenzia europea per la protezione dei diritti fondamentali, Vienna.

(1) Per una panoramica, cfr. M. Rosenfeld e A. Sajo, *Spreading Liberal Constitutionalism: An Inquiry into the Fate of Free Speech Rights in New Democracies*, in *The Migration of Constitutional Ideas*, a cura di S. Choudry, Cambridge, Cambridge University Press, 2007; K. Loewenstein, *Militant Democracy and Fundamental Rights*, in *Am. Polit. Sci. Rev.*, 1937, 3, 417 ss.

(2) K. Loewenstein, *op. cit.*, 417 ss.

di tutela della libertà di espressione costituisce così un momento indispensabile e necessario a cogliere le criticità connesse a ogni disegno che ambisca a una regolazione della materia dei discorsi d'odio e delle *fake news*. Soprattutto, questa lettura, che incrocia due variabili, quella geografica (Europa-Stati Uniti) e quella cronologica (era "pre-digitale" ed era "digitale") consente di apprezzare l'esistenza di possibili anticorpi nel sedime costituzionale di riferimento per ingaggiare una battaglia che, al giorno d'oggi, sembra, grazie alle potenzialità diffusive della rete, esporre a pericolo alcuni capisaldi delle democrazie costituzionali.

Sin dalle origini, con la nascita dello stato liberale e la garanzia delle libertà negative, rinsaldate all'interno delle carte fondamentali, il costituzionalismo europeo ha collocato in una posizione centrale la libertà di manifestazione del pensiero, facendone il caposaldo di ogni società democratica e la cifra distintiva di questo modello. La centralità della libertà di manifestazione del pensiero si deve non soltanto alla portata simbolica di questa libertà, assunta a cifra distintiva di ogni ordinamento democratico, ma anche al suo stretto collegamento con molti dei diritti e delle libertà che ricevono tutela costituzionale. La partecipazione politica, ad esempio, sia in forma attiva che in forma passiva, presuppone il possesso di un bagaglio informativo da parte del cittadino-elettore e, al contempo, postula che il cittadino-candidato non incontri ostacoli nell'esercizio della libertà di parola. Ma molte altre libertà si fondano sul riconoscimento della libertà di espressione nell'ambito di una società democratica. La centralità occupata dalla libertà di manifestazione del pensiero non deve, tuttavia, indurre a ritenere che il campo di applicazione di questo diritto fondamentale non possa essere soggetto a limitazioni o a restrizioni dovute all'esigenza di prevenire abusi, ovvero di bilanciarne l'esercizio con altri diritti, parimenti meritevoli di tutela costituzionale. Sin dalla prima solenne affermazione, che si ritrova nell'art. 11 della Dichiarazione dei diritti dell'uomo e del cittadino del 1789, la libertà di espressione conosce un carattere di cedevolezza intrinseco, che si evince dalla sua stessa formulazione: «La libera manifestazione dei pensieri e delle opinioni è uno dei diritti più preziosi dell'uomo; ogni cittadino può dunque parlare, scrivere, stampare liberamente, salvo a rispondere dell'abuso di questa libertà nei casi determinati dalla Legge». Questa caratteristica non costituisce un tratto isolato, ma rappresenta l'essenza del "dna" europeo sulla libertà di espressione. Un'essenza che ha trovato sublimazione perfetta e del tutto emblematica allorché, con la Convenzione europea dei diritti dell'uomo, gli Stati membri del Consiglio d'Europa si sono muniti di uno strumento comune di protezione dei diritti umani, articolato su un duplice livello: da un lato, l'affermazione solenne di tale libertà; dall'altro, l'enunciazione di una serie di limitazioni che rispondono alle garanzie tipiche dello Stato liberale, quali: la previsione di un fondamento legislativo; il rispetto del criterio di proporzionalità (necessario in una società democratica); la tutela di interessi costituzionalmente rilevanti. A questo paradigma si modellano, dunque, gli Stati europei, che concepiscono per definizione la tutela della libertà di espressione come un diritto suscettibile di logiche di bilanciamento e, talvolta, cedevole al manifestarsi di preminenti esigenze di tutela di diritti diversi. Così è, per esempio, in Italia dove, come è noto, l'art. 21 della Costituzione enuncia il limite esplicito del buon costume (concetto che per sua natura e per lungimiranza dei padri costituenti è permeabile all'evolversi dei tempi). Inoltre, esistono diverse norme nell'ordinamento giuridico (si pensi a quelle che puniscono la diffamazione) che incorporano dei limiti "impliciti" all'esercizio di tale libertà.

2. Disinformazione, misinformazione, fake news: profili definitori

Al fine di comprendere l'effettiva portata dell'impatto dell'IA e dello sviluppo tecnologico digitale, un aspetto preliminare di particolare rilevanza appare essere quello relativo alla stessa terminologia connessa al fenomeno della disinformazione (3). Invero, nonostante sia invalsa sempre più nel linguaggio comune la locuzione inglese facente riferimento alle cosiddette "*fake news*", l'utilizzo di tale espressione è in anni recenti andata sempre più incontro a critiche non solo all'interno del contesto accademico, ma, soprattutto, nell'ambito del *policy-making*.

In tal senso si è espresso, in particolare, l'*High Level Group on fake news and online disinformation* (HLEG), istituito dalla Commissione europea, il quale, nel suo Rapporto finale (4) pubblicato nel 2018, ha rigettato la locuzione "*fake news*". L'HLEG ha argomentato, evidenziando, essenzialmente, due motivi: in primo luogo, si è rilevato come il concetto di "notizie false" rischi di essere riduttivo in quanto esclude tutta una serie di manifestazioni del fenomeno (5); in secondo luogo, l'utilizzo mediatico e popolare della stessa espressione "*fake news*" ha nel tempo condotto a una connotazione fuorviante del fenomeno, soprattutto alla luce dell'uso strumentale che ne

(3) V., in tal senso, P. Dunn, M. Durante, *Disinformazione e Internet*, in *La politica nel mondo digitale*, a cura di G. Giacomini, L. Taddio, Milano-Udine, Mimesis, 2023, 58 ss.

(4) M. de Cock Buning *et al.*, *A multi-dimensional approach to disinformation: Report of the independent High level Group on fake news and online disinformation*, Lussemburgo, Ufficio delle pubblicazioni dell'Unione europea, 2018.

(5) M. de Cock Buning *et al.*, *op. cit.*, 10. Per esempio, la disinformazione consta non solo di notizie completamente false, ma anche di forme di mescolanza tra elementi di informazione più o meno rispondenti al vero; inoltre, determinate fattispecie di disinformazione possono assumere forme diverse da quelle tradizionalmente associate al concetto di "notizia", comprendendo infatti anche casi quali la creazione di *account* o di reti di follower falsi, nonché la fabbricazione o manipolazione di video, il ricorso a pubblicità targettizzate, attività di *trolling* (soprattutto se organizzate e coordinate) e la diffusione di *meme*. Sul punto, v. inoltre M.J. Dupuis, A. Williams, *The Spread of Disinformation on the Web: An Examination of Memes on Social Networking*, Piscataway, IEEE, 2019, 1412 ss. Sul fenomeno dei *meme*, v. G. Marino, *Semiotics of Spreadability*, in *Punctum*, 2015, 1, 43 ss.

è stato fatto da parte di alcune personalità politiche e dei loro sostenitori, al fine di screditare e respingere notizie, informazioni e opinioni non volute (6).

A fronte di ciò, appare preferibile il ricorso al concetto di “disinformazione”, che, secondo l’HLEG, ricomprende tutte quelle forme di manifestazione della libertà che si sostanziano di quelle informazioni false, inaccurate o fuorvianti, sovente artificiosamente create, le quali siano presentate e diffuse con lo scopo precipuo di trarre un profitto di carattere economico, politico o ideologico e/o di provocare un danno a livello pubblico, ivi inclusa l’ingerenza nei processi elettorali e democratici (7). Coerentemente, anche il piano d’azione per la democrazia europea del 2020 definisce il concetto “disinformazione” come «contenuto falso e fuorviante, diffuso con l’intento di ingannare o ottenere un guadagno economico e che può provocare danni pubblici» (8).

Occorre, peraltro, precisare che la nozione di disinformazione, così intesa, include non tanto quelle condotte che l’ordinamento riconosce come inerentemente illecite (si pensi, per esempio, al caso della diffamazione o a quello della calunnia), ma, piuttosto, tutti quei casi ove la produzione e la diffusione di contenuto falso non integra, di per sé, una fattispecie illecita, ma produce, comunque, un potenziale danno a principi e valori democratici. Sono del resto proprio questi ultimi casi a rappresentare la maggiore sfida per il legislatore nazionale e per quello europeo, in quanto impongono la necessità di temperare l’interesse pubblico connesso al corretto funzionamento dello stato democratico con la necessità di non arrecare uno sproporzionato nocumento alla libertà individuale dei consociati, che pongono in essere condotte che l’ordinamento ritiene, di per sé, lecite.

L’elemento della volontarietà della creazione e diffusione di un contenuto falso, così come la volontà di causare un danno o di ottenere un profitto, rappresenta, peraltro, l’elemento di discriminazione tra il fenomeno della disinformazione e quello, connesso ma ben diverso, della “misinformazione”. Quest’ultimo “disordine informativo” (*information disorder*) (9) si caratterizza, infatti, per un elemento di carattere soggettivo, in quanto implica precisamente l’assenza della volontà di diffondere informazioni false e, pertanto, si sostanzia nella diffusione di materiale considerato genuino (10). È importante notare come in internet, nella maggioranza dei casi, fattispecie di disinformazione tendano a essere successivamente ricondivise e diffuse da soggetti differenti da quelli che le hanno originate: lo stesso contenuto, pertanto, rappresenta un caso di disinformazione con riferimento ai soggetti che lo hanno originato o disseminato nella consapevolezza della sua falsità e con la finalità di causare un danno o trarne beneficio. Diversamente, costituisce un esempio di misinformazione quando taluno, appresa la notizia, abbia contribuito a diffonderla nell’erronea convinzione della sua corrispondenza alla verità. Si tratta, come detto, di un aspetto tipico dell’ecosistema internet, ove i contenuti disinformativi hanno la capacità di propagarsi e assumere facilmente connotati virali (11).

Fenomeno ancora differente – ma di minor rilievo ai fini della presente trattazione – è quello della malinformazione, che consta della diffusione di informazioni rispondenti al vero, o comunque basate su elementi fattuali reali, le quali vengono tuttavia comunicate e diffuse al preciso fine di provocare un danno.

Sebbene in misura diversa, i tre menzionati fenomeni – disinformazione, misinformazione e malinformazione – pongono potenzialmente importanti pericoli per il buon funzionamento della società democratica, in quanto impattanti sui processi decisionali della stessa popolazione. In particolare, la disinformazione mira direttamente a determinare una manipolazione o cattiva rappresentazione della realtà, al fine di alterare l’opinione pubblica, ovvero di rafforzare determinate posizioni favorevoli all’agente “disinformatore”: proprio per questo motivo, essa, intesa come sintesi tra l’elemento oggettivo della falsità e l’elemento soggettivo del “dolo” riferibile alla consapevolezza della falsità e volontà della diffusione di tale falsità per finalità di ordine egoistico, ha raccolto su di sé l’attenzione del legislatore sia euro-unitario, sia nazionale.

(6) M. de Cock Buning *et al.*, *op. cit.*, 10.

(7) Così M. de Cock Buning *et al.*, *op. cit.*, 10: «We define it as false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit. The risk of harm includes threats to democratic political processes and values, which can specifically target a variety of sectors, such as health, science, education, finance and more. It is driven by the production and promotion of disinformation for economic gains or for political or ideological goals, but can be exacerbated by how different audiences and communities receive, engage, and amplify disinformation».

(8) Comunicazione COM/2020/790 della Commissione del 3 dicembre 2020 al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni sul piano d’azione per la democrazia europea, 20.

(9) C. Wardle, H. Derashkan, *Information Disorder: Toward an interdisciplinary framework for research and policy making*, Rapporto n. DGI(2017)09 del Consiglio d’Europa, in <www.rm.coe.int>, 27 settembre 2017.

(10) Uno storico esempio di misinformazione, descritto in C. Jack, *Lexicon of Lies: Terms for Problematic Information*, in *Data & Society Research Institute*, 9 agosto 2017, 2, <https://datasociety.net/pubs/oh/DataAndSociety_LexiconofLies.pdf>, fu quello, assai antecedente rispetto all’avvento di internet, in cui il *Chicago Daily Tribune*, basandosi sulle opinioni dei suoi analisti politici e sui sondaggi elettorali dei giorni precedenti, aveva erroneamente riportato nella sua edizione del mattino che Dewey avesse sconfitto Truman nelle elezioni presidenziali statunitensi del 1948.

(11) Così C. Wardle, *Understanding information disorder*, in <www.firstdraftnews.org>, 22 settembre 2020: «When disinformation is shared it often turns into misinformation. Misinformation also describes false content, but the person sharing doesn’t realize that it is false or misleading. Often a piece of disinformation is picked up by someone who doesn’t realize it’s false and that person shares it with their networks, believing that they are helping».

Peraltro, giova sottolineare come lo stesso concetto di disinformazione possa includere al suo interno una pluralità di fattispecie differenti, più o meno gravi (12), le quali impongono chiaramente la necessità di adottare strategie e soluzioni differenziate e calibrate, soprattutto alla luce della necessità di garantire nel contempo la piena tutela della libertà di espressione e di informazione. Inoltre, come segnalato dallo stesso HLEG, la disinformazione, costituendo essa stessa un fenomeno multiforme, richiede di essere affrontata attraverso il ricorso a un approccio “multi-dimensionale”, capace di includere, oltre a misure di carattere giuridico-sanzionatorio, anche e soprattutto interventi volti alla promozione della trasparenza, allo sviluppo dell’alfabetizzazione mediatica e informativa dei cittadini, all’emancipazione e al potenziamento di utenti e giornalisti, nonché alla promozione di un sistema mediatico pluralistico, connotato da diversità e sostenibilità (13).

3. Intelligenza artificiale e produzione della disinformazione

I sistemi di IA possono ricoprire una pluralità di ruoli differenti nel ciclo di vita della disinformazione, investendo sia la produzione e la creazione, sia la disseminazione; inoltre, la stessa IA rappresenta essa stessa un potenziale strumento utilizzabile nel contrasto e nella lotta contro il fenomeno della disinformazione.

Sotto il primo profilo, quello relativo alla produzione della disinformazione, va sottolineata l’emersione di innumerevoli sistemi, legati in particolare alla cosiddetta “IA generativa”, ai modelli fondativi e ai *large language models*, capaci di creare immagini, video e testi sintetici altamente realistici. In tal senso, il numero di esempi concernenti l’utilizzo di sistemi di IA per la produzione di contenuti falsi (e in particolare *deepfake*) è particolarmente allarmante nel contesto odierno, soprattutto alla luce dello scoppio di conflitti – quello russo-ucraino e quello israeliano-palestinese – il cui impatto sullo scenario geopolitico mondiale appare essere determinante.

A puro titolo indicativo del crescente ruolo dell’IA nella produzione di disinformazione e manipolazione dell’informazione, appare di particolare interesse un dato che emerge dalla banca dati di *NewsGuard*, dedicata alla raccolta e catalogazione di narrative false (e virali) in rete. Nel periodo tra l’1 gennaio 2021 e il 30 maggio 2024, il database contiene 930 voci di false narrative: di queste 930 voci, 51 (ovverosia, il 5,48%) sono catalogate come notizie generate attraverso l’uso dell’IA (*AI-Generated Media*) o, comunque, come media manipolati o artificialmente fabbricati (*Manipulated or Fabricated Media*). È, tuttavia, da notare come la percentuale di tali categorie di contenuti generati attraverso l’uso di IA (o comunque manipolati, o fabbricati) sulla quantità delle “bufale” raccolte, sia andata crescendo in modo alquanto significativo nel (breve) lasso di tempo considerato: 1 su 127 (0,78%) nel 2021; 6 su 253 (2,37%) nel 2022; 22 su 293 (7,5%) nel 2023; 22 su 257 (8,5%) nella prima parte del 2024 (14). Pur nella consapevolezza delle limitazioni di una tale ridotta ricerca empirica, si ritiene che tali dati offrano quanto meno un affaccio sull’importante incremento dell’uso dell’IA nel contesto della disinformazione.

È altrettanto rilevante, ai fini di una più chiara illustrazione della questione, proporre alcuni esempi particolarmente significativi di tali contenuti creati o manipolati ad arte attraverso l’uso di sistemi di IA.

A novembre 2023, negli Stati Uniti, un video veniva fatto circolare in internet in cui figurava la rappresentante democratica Alexandra Ocasio-Cortez intenta a discutere in modo (a dir poco) sconclusionato il tema della richiesta di un cessate-il-fuoco a Gaza: «*Cease-fire means that somebody sees a fire. It could be any kind of fire. It could be a big fire or a small fire, a bonfire or even a candle flame. It just matters that somebody sees a fire – that’s why we call it a cease-fire*». Il video, riferibile, attraverso un *watermark* visibile sullo sfondo, a uno specifico account di X (C3PMeme), era, chiaramente, un *deepfake* ottenuto attraverso l’alterazione di un estratto di una *live session* su Instagram, ove Ocasio-Cortez aveva espresso le sue posizioni favorevoli a un cessate-il-fuoco (15). Peraltro, nonostante la chiara falsità del prodotto audiovisivo, esso riceveva una significativa eco non solo su X, ma anche su una pluralità di *social network* differenti, raggiungendo, pertanto, un’ampia fetta della popolazione. Se, da un lato, la chiara insensatezza del discorso poteva apparire sotto certi profili ai limiti del comico, dall’altro lato è chiaro come la circolazione di un simile falso avesse il potenziale di minare in modo significativo la stessa autorevolezza di un personaggio politico come Ocasio-Cortez e, pertanto, in modo più o meno diretto, di alterare la percezione politica della popolazione stessa.

Nello stesso periodo, veniva diffusa una registrazione contenente un messaggio audio in cui, apparentemente, il sindaco di Londra Sadiq Khan avrebbe dichiarato la volontà di posticipare l’*Armistice Day*, giornata dedicata nel Regno Unito al ricordo dei tragici eventi della Prima guerra mondiale, per consentire l’organizzazione di una marcia in favore della Palestina (16). L’audio, originariamente diffusosi su TikTok, si era successivamente

(12) C. Wardle, *Fake news. It’s complicated, First Draft*, in <www.medium.com>, 16 febbraio 2017.

(13) M. de Cock Buning et al., *op. cit.*, 10.

(14) Dati tratti dal database *Misinformation Fingerprints* di *NewsGuard*. V. <<https://www.newsguardtech.com/it/solutions/misinformation-fingerprints>>.

(15) Reuters Fact Check, *Fact Check: Video of Ocasio-Cortez explaining ceasefires is digitally altered*, in <www.reuters.com>, 24 novembre 2023.

(16) «*The Prime Minister meeting with Mark yesterday is a complete waste of time. The buck stops with me. Mark reports to me. I know we have Armistice Day on Saturday, but why should Londoners cancel the Palestinian march on Saturday? Why don’t they have Remembrance weekend next weekend? What’s happening in Gaza is much bigger than this Weekend*

espanso anche su altre piattaforme, venendo amplificato soprattutto da *account* per lo più vicini a idee di destra e indignati dalle supposte dichiarazioni di Khan.

L'IA si presta, in sostanza, a una pluralità di possibili utilizzi finalizzati alla produzione di materiali e contenuti aventi natura eminentemente disinformativa (17) – e, talora, come accade anche nel caso delle cosiddette “allucinazioni” (18), anche aventi natura “meramente” misinformativa.

Come posto in luce da un recente studio pubblicato dallo European Digital Media Observatory (EDMO) (19), le tecniche più diffuse allo stato dell'arte, con riferimento alla produzione di immagini e video *deepfake*, includono, tra le altre, le seguenti:

- la manipolazione di attributi facciali (*face attribute manipulation*), attraverso cui vengono modificate certe caratteristiche riferibili al volto della persona ritratta (per esempio, invecchiamento o ringiovanimento del viso);
- tecniche di scambio volti (*face swap*), attraverso le quali al volto presente nell'immagine o video originale se ne sostituisce un altro (tecnica, peraltro, frequentemente utilizzata per la realizzazione di *deepfake* a sfondo sessuale);
- tecniche di *face reenactment* e *lip synching*, attraverso le quali un video viene manipolato in modo tale modificare i movimenti e le espressioni del volto della persona ritratta in modo tale da far sembrare che esse abbiano reso determinate dichiarazioni (è questo, per esempio, il caso sopra menzionato relativo al video di Ocasio-Cortez).

Al tempo stesso, come evidenziato anche dall'esempio concernente il messaggio audio falsamente attribuito a Sadiq Khan, l'industria dell'intelligenza artificiale è andata evolvendosi vertiginosamente anche nel campo della sintesi vocale. In tale contesto, in particolare, sono andati diffondendosi sia, da un lato, sistemi capaci di sintetizzare contenuti audio a partire da *input* testuali (*text-to-speech*) sia, dall'altro lato, sistemi capaci di modificare tali contenuti audio in modo tale da trasferire su di essi le particolarità vocali, incluso il timbro, di una determinata persona (*voice conversion*) (20).

Ancora, come già più volte menzionato, l'IA, soprattutto a seguito dello sviluppo di LLM, trova oggi importanti sfere applicative nella stessa produzione automatizzata di testi: peraltro, recenti studi hanno confermato come molti di tali sistemi, a fronte di *prompt* volti alla produzione di contenuto disinformativo, siano capaci e abbiano la tendenza a creare articoli e notizie, apparentemente realistici ma non realmente corrispondenti alla verità dei fatti, sulla base di tali suggerimenti; solo alcuni sistemi, a fronte di tali *prompt*, tendono, tuttavia, a produrre testi che contraddicano questi ultimi o, comunque, solo in taluni casi sono esistenti filtri di sicurezza volti a ridurre la produzione di disinformazione (21).

Chiaramente, la diffusione e l'uso massivo di tali sistemi implica potenziali rischi di particolare importanza nel contesto dell'espletamento di processi democratici. Tale rischio è stato, in particolare, posto in luce, da ultimo, nelle linee-guida redatte dalla Commissione europea a inizio 2024 – alla luce dell'avvicinarsi delle elezioni europee – finalizzate a fornire indicazioni specifiche ai gestori di piattaforme e motori di ricerca online di grandi dimensioni per il rispetto delle loro obbligazioni relative all'attenuazione di rischi sistemici connessi a «eventuali effetti negativi, attuali o prevedibili, sul dibattito civico e sui processi elettorali, nonché sulla sicurezza pubblica» (22). All'interno di tali linee-guida, la Commissione argomenta specificatamente quanto segue:

and it's current». V. T. Thompson, No evidence clip of Sadiq Khan supposedly calling for 'Remembrance weekend' to be postponed is genuine, in <www.fullfact.org>, 10 novembre 2023.

(17) V., tra gli altri, N. Giansiracusa, *How Algorithms Create and Prevent Fake News: Exploring the Impacts of Social Media, Deepfakes, GPT-3, and More*, Acton, Apress, 2021, 49 ss.

(18) Le “allucinazioni” costituiscono gli esiti errati delle elaborazioni realizzate dall'IA. Alcuni esempi di allucinazioni sono descritti in G. Finocchiaro, *Intelligenza artificiale. Quali regole?*, Bologna, Il Mulino, 2024, 24: «il caso in cui ChatGPT ha creato delle sentenze inesistenti, di pura fantasia, favorevoli all'avvocato che aveva utilizzato il sistema di IA per redigere l'atto; quello in cui ha costruito vere e proprie prove documentali *ad hoc*, diffamando un soggetto; e ancora i casi in cui è stata fabbricata documentazione scientifica inesistente».

(19) K. Bontcheva (a cura di), *Generative AI and Disinformation: Recent Advances, Challenges, and Opportunities*, in <www.edmo.eu>, 13 febbraio 2024, 4 ss.

(20) Così K. Bontcheva (a cura di), *op. cit.*, 6 s.: «Recent advancements have yielded high-quality models that are accessible to non-expert users, offering both text-to-speech and voice conversion: Text-to-speech involves converting written text into spoken words using synthetic voice and voice conversion entails transferring the unique characteristics of a speaker's voice from an audio recording onto another set of spoken words... While speech synthesis has emerged as a powerful tool for disinformation creation, its potential for misuse (e.g. to damage citizen trust or undermine democratic processes) has long been underestimated».

(21) K. Bontcheva (a cura di), *op. cit.*, 9; I. Vykopal et al., *Disinformation Capabilities of Large Language Models*, Ithaca, Cornell University Library, 15 novembre 2023, <https://arxiv.org/abs/2311.08838>.

(22) Regolamento (Ue) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/Ce (regolamento sui servizi digitali, Digital Services Act, DSA), PE/30/2022/REV/1, art. 34(1)(c).

«Recent technological developments in generative AI have enabled the creation and widespread use of artificial intelligence systems capable of generating text, images, videos, or other synthetic content. While such developments may bring many new opportunities, they may lead to specific systemic risks in the context of elections. Notably, generative AI can be abused to mislead voters or to manipulate electoral processes by creating and disseminating inauthentic, biased, misleading synthetic content (including text, audio, still images and video) regarding political actors, false depiction of events, election polls, contexts or narratives. Generative AI systems can also produce incorrect, incoherent, or fabricated information, so called “hallucinations”, that misrepresent reality, and which can potentially mislead voters (23)».

In sostanza, dunque, la Commissione sottolinea che, se l'IA generativa offre importanti nuovi strumenti per la produzione dell'informazione e, quindi, apre le porte a un arricchimento significativo dello stesso panorama informativo, essa impone al tempo stesso di far fronte alle nuove minacce connesse alla manipolazione dei processi democratici.

A fronte di tali rilievi, la Commissione sottolinea l'assoluta importanza di adottare strategie di contrasto alla produzione di contenuti falsi capaci di ingannare e, dunque, di manipolare le idee della popolazione, garantendo peraltro, nel contempo, la tutela degli stessi diritti fondamentali degli utenti (quale per l'appunto, la libertà di espressione e di informazione) tutelati dalla Carta dei diritti fondamentali dell'Unione europea (24). Tali strategie possono includere, tra l'altro, lo sviluppo di tecniche, quale, per esempio, l'apposizione “watermark”, attraverso le quali i contenuti prodotti dall'IA possano essere facilmente individuati dal pubblico (25); o ancora, l'adeguamento dei termini e delle condizioni dei *provider* oltre che l'adattamento delle stesse strategie di applicazione di tali termini e condizioni (26).

Sotto questo profilo è, d'altro canto, intervenuto anche il nuovo Regolamento europeo sull'intelligenza artificiale il quale, pur avendo un campo applicativo che va chiaramente oltre quello relativo al contrasto alla disinformazione e alla tutela di un dibattito democratico pluralistico, ha cionondimeno introdotto un'importante norma destinata a contenere precisamente i rischi derivanti dall'uso di IA generativa e dalla produzione di *deepfake*. L'articolo 50 del nuovo Regolamento, infatti, introduce alcuni “obblighi di trasparenza per i fornitori e gli utenti di determinati sistemi di IA” i quali includono in particolare:

a) l'obbligo per i fornitori di sistemi di IA (compresi i modelli di GPAI), capaci di generare contenuti audio, immagini, video o testuali sintetici, di garantire che tali contenuti siano marcati in un formato leggibile meccanicamente e rilevabili come generati o manipolati artificialmente (27); nonché

b) l'obbligo per i *deployer* di sistemi di IA generativa capace di produrre audio o video costituenti *deepfake* di rendere noto che tali contenuti sono stati generati o manipolati artificialmente e, allo stesso tempo, l'obbligo per i *deployer* di sistemi di IA per la generazione o manipolazione di testi pubblicati allo scopo di informare il pubblico su questioni di interesse pubblico di rendere noto che il testo è stato generato o manipolato artificialmente (28).

Scopo della norma, alla quale si accompagna, tra l'altro, ricalcando il modello co-regolativo del DSA, la possibilità di elaborare «codici di buone pratiche a livello dell'Unione per facilitare l'efficace attuazione degli obblighi relativi alla rilevazione e all'etichettatura dei contenuti generati o manipolati artificialmente» (29), è tra l'altro proprio quello di ridurre l'impatto che l'IA ha a livello informativo sui cittadini dell'Unione: in tal senso, il legislatore mira, dichiaratamente, a ridurre i «nuovi rischi di cattiva informazione e manipolazione su vasta scala» (30).

4. Intelligenza artificiale e diffusione della disinformazione

Oltre che per la produzione di contenuti di disinformazione, l'IA ricopre un ruolo fondamentale nel contesto della (maggiore o minore) diffusione di tali contenuti. Con riferimento a tale aspetto, almeno due profili sembrano essere di particolare rilievo: da un lato, l'utilizzo di IA da parte dei produttori o di soggetti comunque interessati a promuovere la disseminazione di materiali falsi in rete al fine precipuo di aumentarne l'impatto; dall'altro lato, il peculiare sistema di organizzazione dei contenuti in internet.

(23) Comunicazione della Commissione-Orientamenti della Commissione per i fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi sull'attenuazione dei rischi sistemici per i processi elettorali a norma dell'articolo 35, paragrafo 3, del regolamento (Ue) 2022/2065, par. 36 <<https://digital-strategy.ec.europa.eu/it/library/guidelines-providers-vlops-and-vloses-mitigation-systemic-risks-electoral-processes>>.

(24) DSA, par. 38.

(25) DSA, par. 39(a).

(26) DSA, par. 40(a).

(27) Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (AI Act, AIA), COM/2021/206 final, art. 50(2).

(28) AIA, art. 50(4).

(29) AIA., art. 50(7).

(30) AIA., considerando 133.

Per quanto concerne il primo profilo, particolarmente diffusa è, infatti, la pratica di ricorrere a *social bot*, ovverosia ad account falsi gestiti in modo automatico o semiautomatico (in quest'ultimo caso si parla di "cyborg", ovverosia di profili gestiti in parte da persone vere e in parte dall'IA) con il preciso obiettivo di contribuire alla diffusione di materiali "inquinanti" (non solo disinformazione ma anche, per esempio, *hate speech*) (31). Proprio per far fronte a tale problema, il citato articolo 50 dell'AIA prevede, oltre agli obblighi appena descritti, anche l'ulteriore regola che «i sistemi di IA destinati a interagire direttamente con le persone fisiche siano progettati e sviluppati in modo tale che le persone fisiche interessate siano informate del fatto di stare interagendo con un sistema di IA» (32).

Per quanto concerne il secondo profilo, l'IA veste un ruolo di fondamentale importanza nel contesto della stessa cura dei contenuti operata dalle piattaforme digitali. Di particolare rilievo sono, in questo contesto, i sistemi di raccomandazione (*recommender o recommendation systems*), i quali, partendo dai dati e dalle informazioni raccolte sulle preferenze del singolo utente, sono in grado di predire l'indice di gradimento con riferimento a nuovi contenuti ed elementi (33). Così, per esempio, le piattaforme di *video-sharing*, quali per esempio YouTube o Vimeo, processano le informazioni relative ai video visualizzati in precedenza dall'utente per suggerire allo stesso ulteriori contenuti ad essi simili; allo stesso modo, le piattaforme di programmi *on demand* (Netflix, Amazon Prime, Disney+ ecc.) utilizzano tali sistemi per rendere più facilmente fruibili i loro servizi ai clienti.

I sistemi di raccomandazione costituiscono, peraltro, strumenti essenziali anche nel contesto dei *social network*, permettendo ai fornitori di tali piattaforme di garantire ai loro utenti di essere raggiunti da quei contenuti maggiormente rispondenti non solo ai loro gusti e interessi personali, ma anche al loro sistema di valori e al loro orientamento ideologico-politico. Appare, pertanto, evidente il ruolo centrale che tali sistemi di raccomandazione giocano in generale nella diffusione delle informazioni e, pertanto, nella formazione della coscienza pubblica: si è correttamente osservato che essi hanno la capacità di influenzare e strutturare le preferenze degli utenti e di guidarne le scelte sia a livello individuale, sia a livello sociale e collettivo (34).

L'utilizzo di sistemi automatizzati di cura dei contenuti in rete rappresenta un'arma a doppio taglio nel contesto del contrasto alla disinformazione e della promozione del pluralismo informativo. In effetti, come si specificherà ulteriormente più avanti, i sistemi di raccomandazione possono verosimilmente essere utilizzati con la finalità precipua di arricchire il panorama informativo, favorendone in modo sistemico la natura pluralistica.

Al tempo stesso, sotto il profilo dei possibili effetti collaterali del ricorso ai sistemi di raccomandazione, si è sottolineato da più parti come lo sviluppo di questi ultimi sia in realtà per lo più orientato all'interesse – economico – della massimizzazione dell'*engagement* degli utenti e, pertanto, dei profitti stessi. Poiché, peraltro, contenuti altamente divisivi e polarizzanti – come sono del resto sovente le notizie false (si ripensi, in tal senso, agli esempi relativi al video di Ocasio-Cortez e all'audio attribuito a Sadiq Khan di cui al paragrafo precedente) – tendono frequentemente ad attrarre e catalizzare intorno a sé l'attenzione del pubblico, si pone il significativo rischio che l'algoritmo, pur di suscitare l'interesse degli utenti, sia disincentivato a ridurre la diffusione di disinformazione (se non, addirittura, incoraggiato a promuoverne la visibilità) (35).

Non si intende, in questa sede, affermare che i gestori e fornitori di piattaforme online mirino volontariamente a una maggiore diffusione di tali tipologie di contenuti: al contrario, come si illustrerà più avanti, si è da più parti

(31) V., per esempio, N. Bontridder, Y. Pouillet, *The role of artificial intelligence in disinformation*, in *Data&Policy*, 2021, 3, 32. V., inoltre, M.R. Carbone, *Il lato oscuro dei bot, macchine di disinformazione: come difendersi*, in <www.agendadigitale.eu>, 3 luglio 2020.

(32) AIA, art. 50(1).

(33) V., sul tema dei sistemi di raccomandazione, S. Milano, M. Taddeo e L. Floridi, *Recommender systems and their ethical challenges*, in *AI & Soc*, 2020, 35, 957 ss.

(34) Così S. Milano, M. Taddeo e L. Floridi, *op. cit.*, 957: «After all, [recommender systems] shape user preferences and guide choices, both individually and socially. This impact is significant and deserves ethical scrutiny, not least because [recommender systems] can also be deployed in contexts that are morally loaded, such as health care, lifestyle, insurance and the labour market».

(35) Così E. Llansó et al., *Artificial Intelligence, Content Moderation, and Freedom of Expression*, in *A working paper of the Transatlantic Working Group on Content Moderation Online and Freedom of Expression*, 26 febbraio 2020, 15 ss. <<https://www.ivir.nl/publicaties/download/AI-Llanso-Van-Hoboken-Feb-2020.pdf>>: «These social media platforms rely on advertising revenue and their business model centers on engagement, or keeping users "hooked". Controversial, provocative, and extreme content can drive engagement, so their algorithms learn to prioritize such content when it is popular with users... The effects of recommended content are highly unpredictable. Nevertheless, actors with more oversight, resources, and/or knowledge of platform dynamics are general better able to manipulate the wider information landscape. The threat became visible during the "fake news" controversy around the 2016 U.S. presidential election and UK's Brexit referendum, and the subsequent Cambridge Analytica/Facebook scandal... these stories do underscore the amplification of harmful content through recommendation systems is not necessarily fully intentional or expected on the part of the platform... Wittingly or not, platforms may actively contribute to the amplification of incendiary, controversial and divisive (dis)information as it directly aligns with the commercial and technological infrastructures of their recommendation systems that are optimized for user engagement». Sul tema v. tra gli altri, altresì, G. Chaslot, *How Algorithms Can Learn to Discredit "the Media"*, in <www.guillaumechaslot.medium.com>, 1 febbraio 2018.

cercato – anche da parte dell’industria di settore – di adottare soluzioni tecniche e di *policy* volte a ridurre simili effetti collaterali. È chiaro, tuttavia, che l’orientamento di tali sistemi a una massimizzazione dell’*engagement* determini il potenziale rischio di una predilezione di tali contenuti: una predilezione che, pur non voluta direttamente dall’agente umano, ben può derivare dall’autonoma elaborazione dei dati da parte del sistema di IA.

Tale rischio è particolarmente acuto alla luce del fenomeno delle camere dell’eco e delle bolle-filtro, in quanto – soprattutto all’interno di gruppi sociali legati a teorie cospirazionistiche – l’interazione con contenuti di disinformazione apre alla concreta possibilità di ulteriore esposizione ad altra disinformazione (36). Inoltre, numerosi studi hanno posto in luce come l’IA applicata ai sistemi di moderazione e cura dei contenuti sia essa stessa esposta a errori e *bias* a danno, soprattutto, di minoranze, ovvero di categorie vulnerabili o discriminate, con significativi impatti sull’effettiva promozione di un ecosistema informativo pienamente pluralistico (37).

La necessità di tenere in considerazione il ruolo preminente dei sistemi di raccomandazione, soprattutto ai fini del contrasto alla disinformazione, è stato del resto posto in luce dallo stesso DSA, il quale, con riferimento ancora una volta ai meccanismi di valutazione e attenuazione dei rischi sistemici di cui agli articoli 34-35, sottolinea il dovere dei fornitori di piattaforme e motori di ricerca online di grandi dimensioni di «concentrarsi sui sistemi o su altri elementi che possano contribuire ai rischi, compresi tutti i sistemi algoritmici che possano essere pertinenti, in particolare i loro sistemi di raccomandazione e i loro sistemi pubblicitari» e, precisamente in questa ottica, di «prestare particolare attenzione al modo in cui i loro servizi sono utilizzati per diffondere o amplificare contenuti fuorvianti o ingannevoli, compresa la disinformazione» (38). Per di più, il DSA specifica che, tra le altre misure potenzialmente adottabili, i fornitori di tali servizi «potrebbero dover attenuare gli effetti negativi delle raccomandazioni personalizzate e correggere i criteri utilizzati nelle loro raccomandazioni» (39).

Coerentemente, la Commissione, all’interno delle linee-guida relative all’attenuazione dei rischi sistemici concernenti il regolare svolgimento dei processi democratici, sottolinea l’impatto significativo che tali sistemi di raccomandazione hanno nella formazione del panorama informativo e, di conseguenza, nell’orientamento della volontà collettiva (40). La Commissione suggerisce pertanto di adottare specifiche misure in tal senso e, in particolare, di:

- assicurare che i sistemi di raccomandazione siano programmati e adattati in maniera tale da garantire agli utenti effettive possibilità di scelta e controllo sui loro *feed*, avendo riguardo della diversità e del pluralismo mediatico;
- stabilire misure per ridurre la diffusività della disinformazione nel contesto di elezioni sulla base di metodologie chiare e trasparenti (con riferimento, per esempio, a contenuti ingannevoli che siano stati individuati come falsi attraverso metodologie di *fact-checking* o che provengano da profili cui sia stata ripetutamente ascritta la diffusione di disinformazione);
- stabilire misure per limitare la diffusione, attraverso i loro sistemi di raccomandazione, di contenuti ingannevoli, falsi o fuorvianti, generati con l’ausilio di IA nel contesto di elezioni;
- valutare regolarmente la *performance* e l’impatto dei sistemi di raccomandazione e di affrontare qualsiasi rischio emergente o problema relativo ai processi elettorali, anche attraverso l’aggiornamento e l’affinamento di *policy*, pratiche e algoritmi;
- stabilire misure per garantire la trasparenza del *design* e del funzionamento dei sistemi di raccomandazione, soprattutto in relazione ai dati e alle informazioni utilizzate per sviluppare sistemi che promuovano il pluralismo mediatico e la diversità dei contenuti, in modo da facilitare la valutazione e la ricerca da parte di terze parti;
- collaborare con enti esterni per condurre test e simulazioni (in particolare, *adversarial testing* e *red team testing*) per identificare i potenziali rischi dei sistemi di raccomandazione, quali rischi connessi a *bias*, alla

(36) Così R. Montasari, *Cyberspace, Cyberterrorism and the International Security in the Fourth Industrial Revolution: Threats, Assessment and Responses*, Cham, Springer, 2024, 232: «[Social media platforms] are likely to reinforce this behaviour by recommending more disinformation to users who have already interacted with such content, leading to the potential entrapment of individuals within disinformation filter bubbles. These filter bubbles represent distinct information universes that surround individuals with ideas that resonate with their pre-existing beliefs, including false notions and manipulated perceptions of reality, thereby diminishing informational diversity».

(37) Così, per esempio, C. Are, *The Shadowban Cycle: an autoethnography of pole dancing, nudity and censorship on Instagram*, in *Feminist Media Studies*, 2021, 22, n. 8, 2002 ss. analizza i potenziali effetti che gli algoritmi utilizzati dalle piattaforme hanno nei confronti di contenuti ritraenti nello specifico i corpi delle donne. V., inoltre, tra gli altri, J. Guynn, *Facebook while black: Users call it ‘Zucked,’ say talking about racism is censored as hate speech*, in <www.eu.usatoday.com>, 24 aprile 2019; S. Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York, NYU Press, 2018.

(38) DSA, considerando 84

(39) DSA, considerando 88.

(40) Così all’Allegato alla Comunicazione alla Commissione C(2024) 2121 del 26 marzo 2024, *op. cit.*, par. 27(d): «Recommender systems can play a significant role in shaping the information landscape and public opinion».

soggezione a forme di manipolazione o all'amplificazione di disinformazione, disinformazione, di manipolazioni o interferenze informative di origine straniera o di altri contenuti dannosi (41).

Anche il Codice rafforzato di buone pratiche sulla disinformazione del 2022 riconosce il significativo impatto che i sistemi di raccomandazione hanno sulla dieta informazionale degli utenti e, pertanto, riconosce che tali sistemi debbano essere trasparenti e offrire loro la possibilità di scegliere e modificare in ogni tempo le impostazioni relative a come i contenuti vengono loro presentati (42). Per tale motivo, il Codice prevede una serie di impegni assunti dai firmatari i quali includono, tra l'altro, la minimizzazione dei rischi di propagazione virale della disinformazione attraverso l'adozione di pratiche sicure di sviluppo degli algoritmi e l'adeguamento delle proprie *policy* (43), nonché la predisposizione di funzionalità e strumenti a favore degli utenti stessi finalizzati a consentire loro una maggiore autonomia decisionale.

Ancora, il recente *European Media Freedom Act* (44) ha similmente previsto in capo agli utenti un diritto alla personalizzazione dell'offerta dei media, ovvero sia il diritto, attinente precisamente la sfera della cura dei contenuti, di modificare facilmente la configurazione di qualsiasi dispositivo o interfaccia che controlli o gestisca l'accesso ai servizi di media che forniscono programmi e il loro utilizzo, al fine di «personalizzare l'offerta di media in linea ai loro interessi o preferenze in applicazione del diritto dell'Unione» (45).

Il *focus* dell'Unione in quest'area appare dunque essere, in sostanza, incentrato sull'obiettivo di accrescere non solo l'*accountability* delle piattaforme con riferimento alle modalità di presentazione e diffusione della disinformazione, ma anche, e soprattutto, sulla promozione di una maggiore trasparenza concernente il funzionamento dei sistemi di raccomandazione e sul rafforzamento della posizione soggettiva degli individui che quelle piattaforme utilizzano.

5. Conclusioni: quale futuro per il governo della disinformazione?

Il presente contributo ha esplorato la stretta interrelazione tra IA, disinformazione, principi e processi democratici nel contesto contemporaneo. Come si è avuto modo di esporre, l'affermarsi di nuovi e sempre più sofisticati sistemi di IA ha determinato un salto qualitativo e quantitativo per quanto concerne il fenomeno delle comunemente note “*fake news*” incrementando soprattutto le risorse disponibili ai fini della creazione di contenuti falsi, ma particolarmente realistici e capaci di influenzare l'opinione pubblica, nonché i mezzi per diffondere tali contenuti. Se è vero che l'IA offre anche importanti strumenti utili al contrasto della disinformazione, è, tuttavia, chiaro che essa, unita alle dinamiche tipiche dell'informazione in rete, ha aperto la strada a importantissime sfide per la tutela dei valori democratici.

L'impatto della disinformazione nell'ambito della “società algoritmica” è stato particolarmente evidente in anni recenti con riferimento all'influenza che tale disordine informativo ha esercitato nel contesto di alcuni importanti processi democratici a livello europeo e globale. Del resto, una simile influenza ha evidentemente allarmato una fetta significativa degli attori istituzionali, a livello nazionale e sovranazionale, come si è cercato di porre in evidenza nelle pagine che precedono. Particolarmente emblematica sembra essere l'evoluzione della strategia europea di contrasto alla disinformazione che, nella sua fase attuale, ha subito una drastica accelerazione. D'altro canto, se dal lato europeo si è assistito a un accrescersi della risposta legislativa al fenomeno della disinformazione, il quadro giuridico statunitense in tal senso appare attualmente per lo più invariato: al di là dell'Oceano sembra, infatti, prevalere ancora il magistero del Primo Emendamento.

Un ulteriore aspetto merita, peraltro, di essere posto in luce in questa sede. I nuovi rischi inerenti alla disinformazione e all'inquinamento informazionale sono stati riconosciuti non solo dalle istituzioni pubbliche, statali e/o sovranazionali, ma anche dalle stesse “compagnie del digitale”. Molti di tali attori privati hanno negli ultimi tempi preso consapevolezza della necessità di una maggiore attenzione rispetto alle esternalità negative che un uso scorretto delle loro piattaforme possono potenzialmente causare e hanno, coerentemente, scelto di operare *motu proprio* nella direzione di un contrasto più attivo a fenomeni quale la disinformazione online. Si inserisce, per esempio, in tal senso l'annuncio di febbraio 2024 realizzato da alcune compagnie, quali Adobe, Amazon, Google, IBM, Meta, Microsoft, OpenAI, TikTok e X nel contesto della Conferenza sulla sicurezza di Monaco, dell'adozione di un patto mirato esplicitamente alla neutralizzazione della diffusione di contenuti che possano mettere a rischio l'integrità dei processi elettorali. Tale patto include, in particolare, otto impegni (46):

(41) Allegato alla Comunicazione alla Commissione C(2024) 2121 del 26 marzo 2024, *op. cit.*, par. 27(d).

(42) Codice Rafforzato di Buone Pratiche sulla Disinformazione, 16 giugno 2022, parte V, preambolo, lett. (e) <<https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>>.

(43) Codice Rafforzato di Buone Pratiche sulla Disinformazione, 16 giugno 2022, impegno n. 18.

(44) Regolamento (Ue) 2024/1083 del Parlamento europeo e del Consiglio, dell'11 aprile 2024, che istituisce un quadro comune per i servizi di media nell'ambito del mercato interno e che modifica la direttiva 2010/13/UE, PE/4/2024/REV/1 (regolamento europeo sulla libertà dei media, o *European Media Freedom Act*).

(45) Codice Rafforzato di Buone Pratiche sulla Disinformazione, 16 giugno 2022, art. 20(1).

(46) V. F. González, *Il patto delle big tech per salvare le elezioni dalla minaccia dell'AI*, in <www.wired.it>, 20 febbraio 2024.

- 1) sviluppare e implementare tecnologie per mitigare i rischi legati ai contenuti elettorali ingannevoli creati con sistemi di intelligenza artificiale, compresi quelli *open source*;
- 2) valutare i modelli di IA nell'ambito dell'accordo, allo scopo di comprendere i pericoli che possono comportare in relazione alla produzione di contenuti elettorali ingannevoli;
- 3) rilevare la distribuzione di tali materiali sulle proprie piattaforme;
- 4) adottare misure per gestire in modo appropriato le informazioni ingannevoli distribuite sui propri servizi;
- 5) promuovere una capacità di reazione intersettoriale ai contenuti elettorali ingannevoli;
- 6) fornire al pubblico informazioni che rendano trasparenti le azioni di mitigazione intraprese dalle aziende;
- 7) continuare a confrontarsi con accademici e organizzazioni della società civile;
- 8) sostenere gli sforzi per incentivare la sensibilizzazione, l'alfabetizzazione mediatica e la capacità di reazione della società.

Un altro esempio di buona pratica, in questo senso, è rappresentato dalla predisposizione da parte di Meta, in accordo con la Commissione europea, di una piattaforma finalizzata a permettere agli utenti l'accesso rapido a contenuti e informazioni, in tempo reale, relativi alle sopravvenienti elezioni del Parlamento europeo di giugno 2024. Tale piattaforma include, tra l'altro, la raccolta di post su Facebook e Instagram realizzati dai candidati per ogni stato membro dell'Unione, nonché post realizzati da pagine e account che includano specifiche parole-chiave, ma anche post istituzionali e pubblicati dai partiti (47).

Tali iniziative rappresentano chiari indizi della volontà del settore privato di intervenire attivamente nel contrasto alle forme di inquinamento dell'informazione e, in particolare, aprono la strada a importanti possibilità di strategie di cooperazione tra le istituzioni e le aziende private. Una prospettiva che è, certamente, da incoraggiare, seppure alla luce della rinnovata consapevolezza dell'insufficienza di una strategia di (quasi) delega in bianco alle piattaforme, quale quella che l'Unione aveva inizialmente intrapreso con l'adozione – tra l'altro – del Codice di buone pratiche del 2018.

Il futuro della regolazione della disinformazione all'intersezione con l'IA, in ogni caso, sembra destinato a essere influenzato in modo significativo proprio dai risultati delle consultazioni elettorali del 2024. Permane quindi, allo stato attuale, una cifra di imprevedibilità rispetto alla prossima evoluzione delle strategie descritte.

* * *

(47) Aa.Vv., *2024 European Parliamentary Elections | Live Displays*, in <www.help.crowdtangle.com>, giugno 2024.

INTELLIGENZA ARTIFICIALE, ESIMENTE POLITICA E PROGRAMMAZIONE DELLA SPESA

di Elena Tomassini (*)

Sommario: 1. *Introduzione.* – 2. *Nozione, caratteristiche, utilità dell'intelligenza artificiale.* – 3. *L'intelligenza artificiale, i dati e le proiezioni economiche. La peculiarità dei bilanci degli enti territoriali.* – 4. *L'uso dell'intelligenza artificiale e le decisioni degli organi politici.* – 5. *La giurisprudenza contabile sul ruolo della c.d. "esimente politica" e l'applicazione ai casi di utilizzo dell'IA. nella programmazione della spesa, nell'analisi predittiva dei dati di bilancio e nell'eventuale responsabilità politica ed amministrativa.*

1. *Introduzione.*

La diffusione delle risorse previste nel quadro degli investimenti straordinari del Pnrr e del nuovo ciclo di programmazione europea 2021-2027 fa emergere una questione che attiene alla responsabilità degli enti pubblici nella gestione di questi fondi, di cui sono, da un lato, destinatari e, dall'altro, responsabili dell'erogazione, della corretta gestione e dei relativi sistemi di controllo.

Nuovi strumenti di utilizzo dell'intelligenza artificiale, all'interno di processi e attività della pubblica amministrazione, possono aiutare nella gestione delle risorse pubbliche, in un'ottica di maggiore trasparenza ed efficacia. Al di là delle possibilità investigative per intercettare *ex post* eventuali fenomeni fraudolenti, i sistemi di intelligenza artificiale consentono un'analisi rapida e predittiva, nell'ambito degli investimenti attuati dalla pubblica amministrazione, rendendo il meccanismo di gestione della spesa più efficiente ed evitando lo spreco di risorse pubbliche.

Per Intelligenza artificiale o IA si utilizza il concetto unificato del Regolamento (UE) 2024/1689 del Parlamento Europeo del Consiglio, del 13 giugno 2024 (1).

Inoltre, l'utilizzo dell'intelligenza artificiale può assistere le pubblica amministrazione nella complessiva attività di pianificazione strategica, rendendo l'intero sistema dei controlli molto più efficace; è stato, per esempio, sottolineato che è necessario superare l'attuale sistema proprietario, caratterizzato dal fatto che le singole amministrazioni detengono la proprietà dei propri dati, garantendo il processo di interoperabilità attraverso convenzioni o disposizioni legislative.

2. *Nozione, caratteristiche, utilità dell'Intelligenza artificiale.*

Finora abbiamo parlato di "intelligenza artificiale" o IA (in inglese *AI artificial intelligence*) dandone per presupposta la conoscenza nonché l'univocità del significato. In realtà, la locuzione "intelligenza artificiale" presenta significati diversi, riferibili all'uso di molteplici tecnologie. Si userà, dunque, nel prosieguo, una nozione coincidente con quella di *deep learning supervisionato*, posto che l'utilizzo di "reti neurali profonde" rappresenta ad oggi una delle migliori tecnologie per elaborare proiezioni economiche. Le reti neurali profonde sono modelli matematici ispirati al cervello umano e permettono di approssimare qualsiasi tipo di funzione, lineare e non lineare, a condizione di avere a disposizione sufficiente potenza di calcolo e sufficienti dati nei quali la funzione si sia manifestata (intendendo per funzione una formula matematica idonea a descrivere in termini formali un fenomeno naturale). Il *supervised deep learning* rientra nella più ampia categoria del *supervised machine learning* o apprendimento automatico supervisionato: l'algoritmo (2) riceve una serie di dati storici ed un obiettivo da raggiungere per poi perseguirlo autonomamente. Il *data scientist* (3) non ha, quindi, la necessità di conoscere le

(*) E. Tomassini è Consigliere della Corte dei conti.

(1) Che si prefigge di "migliorare il funzionamento del mercato interno istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di intelligenza artificiale (sistemi di IA) nell'Unione, in conformità dei valori dell'Unione, promuovere la diffusione di un'intelligenza artificiale (IA) antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea («Carta»), compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, proteggere contro gli effetti nocivi dei sistemi di IA nell'Unione, nonché promuovere l'innovazione. Il presente regolamento garantisce la libera circolazione transfrontaliera di beni e servizi basati sull'IA, impedendo così agli Stati membri di imporre restrizioni allo sviluppo, alla commercializzazione e all'uso di sistemi di IA, salvo espressa autorizzazione del presente regolamento". Sul punto e diffusamente, V. Tenore, *L'intelligenza artificiale può sostituire un giudice? I rischi della involuzione verso un "cretino digitale"*, Anicia editrice, 2023.

(2) Un algoritmo è la specificazione di una sequenza finita di operazioni (dette anche *istruzioni*) che consente di risolvere tutti i quesiti di una stessa classe o di calcolare il risultato di un'espressione matematica (da *Wikipedia*).

(3) Il *data scientist* studia e interpreta grandi quantità di dati per ricavarne informazioni utili su cui il committente o l'operatore stesso possa basare le proprie azioni strategiche. Attraverso l'elaborazione dei *Big Data* lo "scienziato dei dati" è in grado di rendere comprensibili le informazioni nascoste nei dati, e in ultimo di trasformare i dati in nuova conoscenza e opportunità.

dinamiche del fenomeno oggetto della ricerca, posto che l'algoritmo provvederà da solo a generare un modello che approssimi le dinamiche del fenomeno studiato (4).

3. *L'intelligenza artificiale, i dati e le proiezioni economiche. La peculiarità dei bilanci degli enti territoriali.*

In base a quanto sopra esposto, è, quindi, possibile applicare l'intelligenza artificiale a dati economici per elaborare delle proiezioni potenzialmente migliori rispetto a quelle ottenute con modelli economici tradizionali. L'applicazione dell'IA ai dati economico-finanziari, a differenza di altri, sconta però numerose problematiche dovute alla peculiarità di questi dati.

I dati economico-finanziari, infatti, sono notoriamente “non stazionari con trend stocastici”: le caratteristiche statistiche dei dati mutano nel tempo, con cambiamenti di *trend* non prevedibili e potenzialmente improvvisi. Le relazioni economiche, infatti, rappresentano un esempio di “sistema adattivo complesso”: le dinamiche di un dato sistema economico si evolvono nel tempo in modo difficilmente prevedibile, posto che gli attori, operanti in un contesto locale o internazionale, cambiano il proprio comportamento in base a fattori esterni alla volontà dell'uomo (esempio fenomeni naturali) ovvero di altri attori in grado di influenzare le dinamiche politiche ed economiche sia su piccola che su grande scala. La rete neurale, quindi, approssima una funzione non stabile nel tempo e, quindi, destinata ad essere superata. La proiezione sconta, quindi, un tasso di errore che sarà tanto maggiore quanto più ampio sarà l'intervallo futuro oggetto della previsione.

Il quadro si complica ulteriormente, nella tematica approfondita, attese sia la natura dei dati della finanza territoriale, sia l'importanza ed il ruolo delle scelte alla base delle decisioni di allocazione delle risorse e delle politiche economiche in generale, che vengono impostate e decise dagli organi politici.

In ordine al primo aspetto, i dati storici dei bilanci scontano una “granularità (5)” ossia il poco frequente aggiornamento; ciò implica la necessità di un intervallo di tempo maggiore intercorrente tra il mutamento delle dinamiche sottostanti e la sua emersione nel dato in modo sufficientemente significativo per poter essere identificato dalla rete neurale; va, altresì, considerato che i dati più recenti acquisiti non possono essere utilizzati per l'addestramento del modello, in quanto necessari per la validazione e test del modello stesso (necessari per confermare che la rete neurale stia effettivamente imparando a generalizzare le dinamiche manifestate nei dati anziché limitarsi a memorizzare i dati storici dando luogo a proiezioni assolutamente inaffidabili). L'elevata latenza tra emersione di una nuova dinamica e sua manifestazione nei dati di *training* della rete neurale rende il modello meno performante, con minore affidabilità delle relative previsioni. La natura “non stazionaria”, dunque, delle dinamiche sottostanti richiede un frequente addestramento del modello, in modo da permettergli di individuare la nuova funzione, ma a ciò si oppone la scarsa granularità dei dati di bilancio.

Sotto il secondo profilo, relevantissimo è il ruolo delle scelte operate a livello politico nazionale e regionale che mutano – assai frequentemente – il contesto normativo, influenzando sulle prospettive economiche di bilancio degli enti territoriali (6). Tali cambiamenti, però, possono essere rilevati dal modello solo quando si manifestano in bilancio, quindi necessariamente in sede di bilancio di previsione, quando tali scelte sono state fatte. Per evitare tale disarmonia, è possibile la creazione di uno o più modelli dedicati di forme di analisi sul dato normativo, per cercare di fornire un dato al modello predittivo, prima che il mutamento si manifesti nei dati di bilancio: il modello di analisi potrebbe assegnare un punteggio numerico alla modifica normativa e tale valore potrebbe essere fornito al modello predittivo, allo scopo di fornire ulteriore contesto per ridurre l'errore delle previsioni.

Attualmente, la necessità di ricorrere a reti neurali estremamente esigenti a livello computazionale, come i *large language models* (Llms) non comporta esigenze di *hardware* talmente costosi da essere proibitivi, potendo basarsi su sistemi “open source” che possono essere personalizzati con costi più limitati (intorno a un milione di euro, comunque). Trattandosi di compiti sensibili idonei ad influenzare l'operato dell'amministrazione, è assolutamente sconsigliabile il ricorso al *cloud computing*, posto che in tale ipotesi l'*hardware* non è nella completa disponibilità dell'amministrazione ma del fornitore e, di conseguenza, si introduce un rilevante vettore di attacco con compromissione dei profili di sicurezza informatica. La necessità di operare *in house* richiede tuttavia notevoli investimenti iniziali ed elevati costi di utilizzo, dato il consumo rilevante di energia da parte dell'*hardware* usato per l'addestramento e per l'inferenza di *large language models*.

(4) Il citato Regolamento europeo, nel tentativo di unificare la nozione di IA, fonda la differenza tra questa e i sistemi automatizzati sulla “capacità inferenziale” dell'IA e principali caratteristiche dei sistemi di IA, che la distinguono dai tradizionali sistemi *software*. Tale capacità inferenziale permette di ottenere un “*output*” diverso dall'apprendimento automatico tramite un approccio basato sulla logica e sulla conoscenza pregressa, consentendo l'apprendimento, il ragionamento o la modellizzazione.

(5) Il concetto di “granularità”, nell'ambito informatico, si riferisce a quanto è divisibile un sistema. Una maggiore granularità rende i dati più flessibili consentendo l'elaborazione separata di parti più specifiche dei dati, ma richiede maggiori risorse di calcolo.

(6) Si pensi ai vincoli di finanza pubblica e ai contributi imposti alle Regioni e ai Comuni dalle leggi di bilancio.

4. *L'uso dell'intelligenza artificiale e le decisioni degli organi politici.*

In base a quanto accennato, emerge quindi come l'uso dell'intelligenza artificiale sia uno strumento ulteriore dato all'organo politico per prendere una decisione; pur con i rilevanti limiti sopra delineati, infatti, e tenuto conto dell'attuale situazione della finanza nazionale e territoriale, il ricorso a tale innovativo strumento può intuitivamente costituire un ausilio formidabile per intercettare le scelte più razionali, efficienti ed efficaci sia in materia di entrata che di spesa, con funzione non soltanto in termini di certezza e veridicità dei dati di bilancio ma e soprattutto in chiave predittiva e di impostazione delle problematiche.

Naturalmente, come sempre avviene quando l'uomo si affida alla "macchina", e tenuto conto della peculiarità del concetto, che mal si attaglia alla IA, gli organi politici, come pure quelli tecnici, non sono certamente esonerati dall'obbligo di operare una scelta autonoma che non si affidi in modo acritico alla previsione del modello di intelligenza artificiale. L'attività predittiva va contestualizzata, infatti, tenendo conto dei limiti della tecnologia e del contesto estremamente mutevole e non stazionario.

A tali complicazioni si aggiunge la natura delle reti neurali, definite "*black box*": sistemi difficilmente intelligibili, anche per il personale tecnico. In altri termini, non è possibile, nemmeno a posteriori, spiegare come mai in concreto una determinata rete neurale abbia restituito un determinato *output* ed è dato solo studiare le *performance* del modello su dati non oggetto di addestramento (dati di validazione e di *test*).

5. *La giurisprudenza contabile sul ruolo della c.d. "esimente politica" e l'applicazione ai casi di utilizzo dell'IA nella programmazione della spesa, nell'analisi predittiva dei dati di bilancio e nell'eventuale responsabilità politica ed amministrativa.*

Com'è noto, la l. 14 gennaio 1994, n. 20, disposizioni in materia di giurisdizione e controllo della Corte dei conti, come modificata dal d.l. 23 ottobre 1996, n. 543, convertito con modificazioni dalla l. 20 dicembre 1996, n. 639, prevede, al c. 1-ter, che "Nel caso di atti che rientrano nella competenza propria degli uffici tecnici o amministrativi la responsabilità non si estende ai titolari degli organi politici che in buona fede li abbiano approvati ovvero ne abbiano autorizzato o consentito l'esecuzione". La giurisprudenza contabile ha dato un'interpretazione restrittiva della portata della norma, statuendo che l'esimente c.d. politica della responsabilità amministrativo-contabile opera soltanto quando la decisione che si assume fonte di danno ingiusto sia stata assunta in materie di particolare difficoltà tecnica o giuridica, dovendosi altrimenti ritenere che l'evidenza dell'erroneità dell'atto sia tale da escludere la stessa buona fede dei titolari dell'organo politico (*ex pluribus*, Corte conti, Sez. I, n. 282/2002). In tale contesto, sono state escluse dal perimetro di applicazione dell'esimente tutti quegli atti emessi dagli organi politici che non riguardano materie di particolare complessità tecnica, ma afferiscono ad attività proprie degli amministratori (esempio richieste di proroga dei termini di realizzazione dei lavori, richieste di erogazione di finanziamento, riscontro ad atti giudiziari, atti che, proprio perché espressione di principi di carattere generale afferenti alle varie fasi del procedimento amministrativo, non rientrano nella competenza propria di un ufficio tecnico, di cui non costituiscono né approvazione né autorizzazione (Corte conti, Sez. giur. Calabria, sent. n. 256/2020). Molte decisioni hanno posto in rilievo l'assorbente responsabilità degli amministratori; il mandato ricevuto dagli elettori affida o impone a sindaci e assessori destinatari delle deleghe, funzioni di indirizzo, vigilanza e controllo non sulla legittimità dei singoli atti di gestione di competenza dell'apparato amministrativo, ma sull'andamento dell'azione amministrativa di attuazione degli indirizzi politici adottati. In tale ambito, è stato altresì precisato che, sebbene la responsabilità del "governo" finanziario dell'ente locale sul versante esecutivo e tecnico sia affidato al ruolo del responsabile del servizio finanziario, quest'ultimo comunque si deve uniformare al principio per cui i poteri di indirizzo e di controllo politico-amministrativo spettano agli organi di governo, che esercitano il potere di decisione politica sulle finanze dell'ente (art. 107 Tuel).

Infatti, nel bilancio confluiscono le scelte politiche e di gestione dell'organo di governo dell'ente, scelte politiche che devono tenere in debito conto le risorse disponibili, nel senso che il bilancio quale bene pubblico è "*funzionale a sintetizzare e rendere certe le scelte dell'ente territoriale, sia in ordine all'acquisizione delle entrate, sia alla individuazione degli interventi attuativi delle politiche pubbliche*" (Corte cost. sent. n. 184/2016).

Tanto premesso, è stata affermata la responsabilità degli organi politici nella redazione dei documenti di bilancio e soprattutto nell'inerzia tenuta nel non aver adottato le misure finanziarie necessarie per recuperare, ad esempio, i disavanzi di amministrazione, emersi a seguito del riaccertamento straordinario dei residui, in spregio al principio di prudenza che costituisce uno dei cardini dei principi in materia di bilancio (economico o finanziario).

Sul punto, relativo non tanto al giudizio sulla corretta e proficua allocazione delle risorse in sede previsionale, ma in termini di recupero del disavanzo emerso in bilancio (risultato di amministrazione) emerge dunque la piena responsabilità degli organi politici; la giurisprudenza ritiene, infatti, prioritaria la funzione di indirizzo e di vigilanza degli amministratori eletti, che, altrimenti, verrebbe svuotata dei contenuti di strumento concreto ed effettivo di esercizio del mandato politico elettorale ricevuto dai cittadini per il perseguimento della migliore gestione degli interessi della comunità territoriale.

È stato, infatti, precisato che la *ratio* dell'esimente in questione risiede nella tutela della buona fede e dell'affidamento degli organi di governo in sede di adozione di atti deliberativi che necessitano di un'istruttoria tecnico-

amministrativa particolarmente complessa, e consegue al principio di netta separazione e divieto di interferenza tra funzioni di indirizzo politico spettanti agli amministratori pubblici e attività gestorie di competenza dirigenziale. In tale contesto, la giurisprudenza contabile ha più volte ribadito che l'esimente non è applicabile nelle materie riservate agli organi di governo, nelle quali gli uffici amministrativi e tecnici della struttura abbiano espletato funzioni istruttorie o consultive e comunque di mero supporto strumentale (Sez. II centr app. n. 111/2018; n. 29/A/1999; n. 303/A/2003; Sez. giur. Lazio, n. 2087/2005; Sez. giur. Lombardia, n. 323/2003; Sez. I centr. app. n. 407/2017; Sez. III centr. app. n. 129/2017).

Inoltre, *“per pacifica giurisprudenza, la suddetta esimente non è applicabile nel caso di atti che, come l’approvazione del rendiconto di esercizio, sono assegnati dalla legge alla competenza di un organo politico”* e *“la presenza di pareri favorevoli e/o l’assenza di rilievi di legittimità da parte del Segretario-Direttore generale o dei Revisori dei conti non sono sufficienti a dispensare i componenti dell’organo politico dalle loro eventuali responsabilità”* e che, in definitiva, *“dispensare i componenti dell’organo politico da ogni responsabilità significherebbe praticamente ridurre, contrariamente alla volontà del legislatore, la deliberazione di Consiglio comunale di approvazione del rendiconto a mera ratifica di decisioni assunte da altri”* (Sez. giur. Campania n. 1071/2018 e n. 314/2019).

Venendo, ora, e in prospettiva, all’applicazione dell’intelligenza artificiale alla fase di previsione del bilancio come pure a quella, altrettanto rilevante, di predizione di eventuali disavanzi o dissesti dell’ente (altra prospettiva in corso di esplorazione tramite modelli da parte della Corte dei conti), pare chiaro, almeno a livello teorico, che ove mai si adottassero modelli di *deep learning* ai campi sopra descritti gli organi politici non potrebbero, parimenti, andare esenti da responsabilità; a meno di non dimostrare che gli organi tecnici che hanno elaborato tali modelli abbiano palesemente errato nelle procedure di raccolta e di analisi dei dati ovvero che, posta la correttezza di queste fasi, a causa delle problematiche sopra indicate (mancanza di stazionarietà dei dati, analisi degli stessi in tempi incompatibili con la loro attendibilità) si escludano, *in re ipsa*, la responsabilità sia degli organi tecnici che di quelli politici. Ma, a questo punto, andrebbe considerata l’opportunità di adozione dei modelli di IA da parte dei decisori politici.

Conclusivamente, anche in questa difficile materia e con gli opportuni accorgimenti l’IA rappresenta un formidabile strumento di ausilio di conoscenza al fine di una migliore programmazione e di una scelta più razionale dei decisori politici. Spetterà alla tecnica conciliare le peculiarità della contabilità pubblica al fine di superare le difficoltà cui si è accennato.

* * *

L'ALGORITMO, TRA INTELLIGENZA ARTIFICIALE E DIRITTO AMMINISTRATIVO

di Daniela Anselmi e Federico Smerchinich (*)

Abstract: L'intelligenza artificiale è ormai parte della nostra società e si sta facendo strada anche nel mondo del diritto pubblico dove, attraverso gli algoritmi, si cerca di rendere più efficienti sia i procedimenti sia i processi. In particolare, gli algoritmi vengono intesi come sistemi che attraverso l'acquisizione e l'incrocio di dati sono in grado di generare degli *output* utili ai procedimenti amministrativi e, a volte, a sostituire l'attività stessa dei dipendenti pubblici. Tuttavia, questo sistema di algoritmi, per quanto idoneo a velocizzare le attività, non può essere lasciato libero da controlli, ma l'uomo deve pur sempre poter verificare il metodo seguito dall'algoritmo stesso e i dati prodotti, al fine di evitare errori o dispersioni di dati sensibili. Proprio le potenzialità in termini di efficienza dell'uso degli algoritmi hanno portato a paventare ipotesi di utilizzo degli algoritmi anche nell'ambito del processo amministrativo in ottica di giustizia predittiva, senza tuttavia sfociare mai in alcuna proposta effettiva dati i molteplici problemi, difficilmente superabili, che riguardano questo tema. Diversamente, nell'ambito dei contratti pubblici il legislatore ha previsto una disciplina specifica nell'art. 30 d.lgs. n. 36/2023 proprio per automatizzare i procedimenti in ottica di efficienza. Nel formulare questa norma, il legislatore ha posto particolare attenzione sul rispetto e sulla garanzia del principio di trasparenza e non discriminazione, oltre che sul ruolo dell'uomo nel poter accedere ai codici sorgente dell'algoritmo e nel poter verificare sia gli *input* forniti all'algoritmo sia gli *output* generati. In una materia in continua evoluzione, sarà fondamentale il coordinamento tra giuristi e tecnici per comprendere tutti gli aspetti dei nuovi sistemi di intelligenza artificiale applicati al diritto.

The artificial intelligence is now part of our society and it is making its way into the field of public law, where algorithms are being used to make both procedures and processes more efficient. In particular, algorithms are considered as systems that, through the acquisition and cross-referencing of data, can generate useful outputs for administrative proceedings and, sometimes, even replace the activities of public employees. However, this system of algorithms, while capable of speeding up activities, cannot be left unchecked; humans must still be able to verify the methods used by the algorithm and the output to avoid errors or the loss of sensitive information. The potential and efficiency offered by the use of algorithms has led to speculation about their application in administrative processes with a view toward predictive justice, although no concrete proposals have emerged due to the numerous issues surrounding this topic. On the other hand, in the area of public contracts, the legislator has established specific regulations in article 30 of d.lgs. n. 36/2023 aimed at automating procedures with a focus on efficiency. In formulating this regulation, the legislator has placed particular attention to ensuring compliance with the principles of transparency and non-discrimination, as well as the role of humans in accessing the source codes of the algorithms and verifying both the inputs provided to the algorithm and the outputs generated. In a constantly evolving field, coordination between legal experts and technical professionals will be essential to understand all aspects of new artificial intelligence systems applied to law.

L'intelligenza artificiale applicata al diritto pubblico, così come al diritto in generale e alla nostra società, è una delle nuove frontiere dell'epoca storica attuale. Per poter correttamente fare ricorso all'intelligenza artificiale, bisogna comprendere che cosa si intende con essa nel mondo giuridico al fine di limitarne il campo ai fini del presente scritto. In particolare, l'intelligenza artificiale nel diritto amministrativo riguarda tutti quei procedimenti di automazione tecnologica che rendono più veloce l'attività umana, finanche a sostituirla in determinati casi.

Questi procedimenti si basano sullo sviluppo di *software* che, tramite algoritmi, consentono di incrociare dati e trovare soluzioni in un tempo esponenzialmente minore rispetto a quello impiegato dall'uomo per compiere la medesima attività.

Stando alla dottrina, quando si parla di algoritmo si intende “una sequenza di passaggi elementari, secondo una sequenza finita e ordinata di istruzioni chiare e univoche (ognuna delle quali eseguibile entro un tempo finito e che produce un risultato in un tempo finito) per la risoluzione di un dato problema” (1). In giurisprudenza è stato affermato che con l'algoritmo “ci si richiama, semplicemente, a una sequenza finita di istruzioni, ben definite

(*) D. Anselmi è avvocato amministrativista presso Studio legale Anselmi&Associati, Genova-Roma, Vicepresidente Unaa, Presidente associazione avvocati amministrativisti Liguri Aaal, Presidente Aidda Delegazione Liguria; F. Smerchinich è avvocato amministrativista dello Studio legale Anselmi&Associati.

(1) V. Neri, *Diritto amministrativo e intelligenza artificiale: un amore possibile*, in *Urb. e appalti*, 2021, 5, 581 ss.

e non ambigue, così da poter essere eseguite meccanicamente e tali da produrre un determinato risultato (come risolvere un problema oppure eseguire un calcolo e, nel caso di specie, trattare un'aritmia)" (2).

Quindi, si tratta di un sistema in grado di incrociare dati e informazioni di *input* ad una velocità estremamente veloce, in grado di fornire un *output* utilizzabile.

Già letta in questi termini la materia potrebbe far sorgere dei dubbi, soprattutto se immaginiamo di utilizzare degli algoritmi all'interno dei procedimenti amministrativi.

Infatti, laddove trattiamo di *input* e *output*, numerose sono le domande che sovengono all'operatore del diritto: Chi inserisce gli *input*? Quali sono gli *input* corretti? Gli *input* inseriti sono completi o parziali? Gli *input* utilizzati proteggono i dati sensibili? L'*output* che mi viene generato è affidabile? Posso utilizzare l'*output* in ogni procedimento amministrativo? L'*output* generato tutela il diritto alla riservatezza?

Molte altre potrebbero essere le domande da porsi, perché il problema che sta alla base degli algoritmi applicati al diritto è legato alla fiducia che il giurista pone nell'algoritmo stesso e al ruolo che l'algoritmo può avere nella vita dell'avvocato, del giudice, del funzionario amministrativo.

Un problema acuito dal fatto che i dati di *input* vengono "donati" al sistema automatico con il rischio di perderne il controllo, soprattutto quando si tratta di dati sensibili.

Problemi che si fronteggiano con l'inevitabile situazione che, attualmente, l'algoritmo e l'intelligenza artificiale appaiono utili (e forse necessari) al giurista, perché in un mondo sempre più regolato in termini di efficienza (data dal coordinamento tra risultato e velocità) è più proficuo utilizzare sistemi che risolvono incognite e trovano soluzioni nel giro di millesimi di secondo.

Proprio la ricerca di questa efficienza, accompagnata dallo sviluppo della tecnologia hardware e dei *software* stanno velocemente portando alla sostituzione di intere fasi del procedimento amministrativo e, forse in futuro, anche di fasi del processo amministrativo.

Infatti, attualmente, nel processo amministrativo l'unica parte che viene regolata dall'intelligenza artificiale è quella che attiene al deposito degli atti in giudizio, dove, attraverso moduli da compilare e allegazioni, è possibile trasmettere a Tar e Consiglio di Stato dei pdf che si portano appresso tutti gli atti e documenti di interesse (nei limiti di 30 o 50 Mb a seconda del sistema).

In tale attività di deposito, il sistema della giustizia amministrativa è già impostato per riscontrare eventuali errori e, se del caso, respingere il deposito, dimostrando che in parte viene già svolta attività di intelligenza artificiale in grado di incidere sul processo stesso (se il deposito viene respinto e scadono i termini per farlo, la parte è tardiva). Un sistema di rilevamento degli "errori bloccanti", che verrà confermato anche nell'emanando decreto di modifica del decreto del Presidente del Consiglio di Stato del 28 luglio 2021, utile ad evitare di incorrere in problemi di inammissibilità dei depositi degli atti amministrativi.

Diversa risulta la tematica dell'intelligenza predittiva applicata al processo, che nel nostro Paese ancora non viene applicata stante i possibili errori che ne potrebbero scaturire. Basti pensare al fatto che, se si sbaglia l'inserimento degli *input*, quindi, fornendo ad esempio alla macchina tutti precedenti giurisprudenziali di uno stesso orientamento, necessariamente la macchina stessa darà un risultato in linea con tali orientamenti, senza tenere conto delle tesi diverse. Un ulteriore ostacolo è dato dal fatto che gli *input* necessari alla giustizia predittiva risultano dalle sentenze dei precedenti giudizi. Sentenze che, spesso, hanno richiami alle parti, a loro situazioni e a dati personali che, inseriti in un algoritmo, rischierebbero di venire diffusi se non correttamente trattati.

Allo stesso modo risulta difficile poter immaginare che una macchina possa comprendere correttamente il contenuto di un ricorso o di una memoria amministrativa tanto da poterne predire l'esito in sentenza.

Per completezza, si coglie l'occasione per rappresentare al lettore che, benché allo stato ancora non risulti esservi alcuna previsione in materia, alcuni hanno paventato l'ipotesi che siano introdotte regole sull'utilizzo di format per la redazione degli atti di parte, al fine di standardizzare la forma o intere parti (ad esempio epigrafe e fatto) degli atti stessi e allenare l'algoritmo in vista di una possibile predizione dei risultati.

Tuttavia, tale sistema è osteggiato dall'avvocatura amministrativa, non solo perché tale standardizzazione potrebbe ledere il diritto di difesa, ma anche perché i giudizi amministrativi si giocano in gran parte su come ogni singolo avvocato redige proprio la parte in fatto, presupposto per l'applicazione del diritto. Inoltre, non pare essere corretto un sistema di "allenamento" dell'algoritmo basato su format standard di atti, basati su modelli non condivisi da tutta la platea degli operatori giuridici.

Così, come risulta non facilmente superabile la tematica dei prezzi degli algoritmi per la spesa pubblica. Difatti, se da una parte colossi come Apple o Google hanno le risorse per sviluppare algoritmi affidabili, difficilmente potrebbe essere immaginabile che le pubbliche amministrazioni possano dotarsi di sistemi altrettanto efficienti.

(2) V. Cons. Stato, Sez. III, 25 novembre 2021, n. 7981; in materia anche Cons. Stato, Sez. VI, 4 febbraio 2020 n. 881; Tar Lombardia, Milano, Sez. II, 31 marzo 2021 n. 843; Cons. Stato, Sez. VI, 8 aprile 2019, n. 2270.

Quindi, per il momento, la giustizia predittiva in senso stretto rimane ancora una tematica all'orizzonte.

Se guardiamo, invece, al lato del procedimento amministrativo, ormai sono alcuni anni che intere attività valutative sono affidate all'intelligenza artificiale anche dal lato sostanziale, cioè non solo quello di acquisizione degli *input* ma anche di generazione degli *output*. Se pensiamo al mondo dei concorsi (principalmente scolastici), la tendenza è spostata verso la velocizzazione delle attività di correzione utilizzando le macchine che possono immediatamente comprendere l'esattezza o meno delle risposte date dal candidato. Tanto che, nel caso delle impugnazioni degli esiti concorsuali (soprattutto scolastici), si va a censurare effettivamente il codice sorgente posto alla base dell'algoritmo, proprio per contestare che i dati di *input* siano errati e, conseguentemente, anche l'*output* dovrebbe esserlo (3).

In tal senso il c.d. "codice sorgente" di un *software* è rappresentato dal testo di un algoritmo di calcolo scritto in un linguaggio di programmazione volto a definire il flusso di esecuzione del programma. Si tratta, in particolar modo, di una sequenza articolata di tutti i dati e di tutti i comandi per mezzo dei quali il programmatore struttura il *software* e ne consente l'esecuzione, determinandone in concreto le modalità di funzionamento (4).

Va da sé, quindi, l'importanza di tale codice, che non ha nulla a che vedere con il diritto, rispetto alla risultanza del procedimento algoritmico stesso che, invece, incide sul mondo giuridico e sugli interessi legittimi dei cittadini.

Addirittura, in alcuni casi, le macchine sono in grado di fornire come *output* il provvedimento finale, mentre in altri coadiuvano i funzionari o il Rup nella raccolta dei dati e delle informazioni necessarie a realizzare la scelta. In altre parole, le macchine possono aiutare gli amministratori nelle diverse fasi procedurali, sostituendosi o affiancandosi ad essi.

In tale scenario il ruolo dell'uomo si sposta dalla corretta esecuzione di tutti i passaggi procedurali, alla legittima scelta del tipo di algoritmo e di macchina da utilizzare per farsi sostituire-aiutare nello svolgere il procedimento.

Un ruolo dell'uomo che rimane, pur sempre, centrale, considerando che non si può arrivare alla spersonalizzazione dei procedimenti amministrativi, ma che è sempre necessario individuare una figura che sovrintenda al lavoro del sistema automatizzato e che ne verifichi le risultanze.

Si parla in tal senso di riserva di umanità nello svolgimento delle funzioni amministrative (5), tanto che in dottrina si auspica un inserimento di questo principio anche nella nostra Costituzione.

Nel settore che cura gli interessi pubblici, dunque, le *machine learning* sono diventate oggetto di un sempre maggiore interesse, inizialmente nei paesi anglosassoni per poi giungere anche in Italia. Ad esempio, negli Stati Uniti ed in Inghilterra, vi sono state le prime regolamentazioni di questi robot, ma successivamente il Parlamento europeo ha approvato una risoluzione sulla disciplina robotica e i sistemi di intelligenza artificiale, confluita nel Regolamento europeo n. 1689/2024 entrato in vigore il 12 luglio 2024. In Italia, invece, l'AgID (Agenzia per l'Italia Digitale) ha pubblicato il Libro bianco sull'Intelligenza Artificiale.

Calandoci nel contesto normativo, bisogna subito rappresentare che è attualmente al vaglio delle Commissioni del Senato il disegno di legge sull'intelligenza artificiale (atto Senato n. 1146 presentato il 20 maggio 2024) che propone un testo generale, volto principalmente a promuovere un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a coglierne le opportunità. Garantisce la vigilanza sui rischi economici e sociali e sull'impatto sui diritti fondamentali dell'intelligenza artificiale.

L'approvazione e l'entrata in vigore di questo disegno di legge non è ancora pronosticabile, ma è certamente un passo avanti nella materia e sarà interessante vederne gli sviluppi.

Questo disegno di legge prevede comunque la necessità di svolgere attività formativa al fine di consentire agli operatori del diritto di poter coscientemente utilizzare gli strumenti dell'intelligenza artificiale.

Guardando alle norme già vigenti, diversi articoli vengono in rilievo quando si applica l'automazione al procedimento amministrativo, potendosi ritenere che l'uso di macchine nell'attività amministrativa possa farsi rientrare all'interno delle scelte per perseguire il buon andamento amministrativo previsto dall'art. 97 Cost. E così, vi è l'art. 3-bis l. n. 241/1990 come modificato dal d.l. n. 76/2020, che lega l'efficienza delle pubbliche amministrazioni all'uso di "strumenti informatici e telematici" sia nell'attività interna all'ente che nel rapporto tra pubblica amministrazione e privati.

Oppure l'art. 50-ter d.lgs. n. 82/2005 che regola i rapporti di interconnessione tra pubblica amministrazione disciplinando anche la Piattaforma digitale nazionale dati.

Benché vi siano delle norme che legittimano l'uso delle risorse automatizzate nel procedimento amministrativo, è comunque importante prestare le giuste cautele in tale attività.

(3) Tar Lazio, Roma, Sez. IV, 22 luglio 2024, n. 14855.

(4) Tar Lazio, Roma, Sez. III-bis, 1 luglio 2020, n. 7526.

(5) G. Gallone, *Riserva di umanità e funzioni amministrative*, Padova, Cedam, 2022.

Difatti, uno dei problemi più concreti che sorgono dalle scelte meccaniche è che non sempre esse forniscono la motivazione che la legge vuole alla base di ogni scelta e provvedimento amministrativo come previsto dall'art. 3 l. n. 241/1990 (6).

Come detto, infatti, le macchine fondano le proprie scelte su algoritmi che fanno della propria esperienza e dei dati acquisiti la base delle scelte future, senza però fornire alcuna giustificazione in merito al percorso logico e giuridico che ha fondato la decisione finale. Ne consegue, che è anche difficile poter sindacare se un algoritmo ha lavorato bene o male e se la scelta finale sia legittima rispetto al caso concreto. Certamente, si possono verificare quali informazioni e dati sono stati posti alla base dell'elaborazione elettronica, ma è complicato comprendere se la scelta abbia seguito dei criteri razionali oppure sia stata indotta da qualche incontrollabile incrocio di dati. Talvolta, è possibile che l'elaborazione di grandi flussi informativi e dati porti alla creazione di modelli che siano inverosimili e lontani dalla realtà, trasformando quello che può essere un aiuto in un ostacolo difficilmente superabile solo dall'algoritmo.

Tuttavia, con il d.lgs. n. 36/2023, l'intelligenza artificiale e l'automazione sono entrate ufficialmente nel mondo della contrattualistica pubblica, al fine di consentire il raggiungimento del superiore principio del risultato nelle gare pubbliche.

E, così, l'art. 30 del nuovo Codice, dal titolo iconico di *“Uso di procedure automatizzate nel ciclo di vita dei contratti pubblici”* prevede al comma 1 che *“1. Per migliorare l'efficienza le stazioni appaltanti e gli enti concedenti provvedono, ove possibile, ad automatizzare le proprie attività ricorrendo a soluzioni tecnologiche, ivi incluse l'intelligenza artificiale e le tecnologie di registri distribuiti, nel rispetto delle specifiche disposizioni in materia”*.

Questo comma correla alcuni principi fondamentali, quali quello dell'efficienza e dell'automatizzazione, come se il primo fosse meglio raggiungibile solo utilizzando il secondo. Non a caso questa norma parte dall'inciso *“per migliorare l'efficienza delle stazioni appaltanti”*. Si fa, altresì, notare che la previsione appena vista è retta da un verbo all'indicativo *“provvedono”*, dimostrando che, salvo deroghe, l'attività amministrativa legata ai contratti pubblici deve (obbligatoriamente) ricorrere a soluzioni tecnologiche, tra cui l'intelligenza artificiale.

Il richiamo della norma al rispetto delle disposizioni in materia si pone come una clausola generale che richiede un necessario coordinamento sia con le altre normative nazionali (soprattutto in tema di dati personali) e con quelle europee, tra cui, da ultimo, il regolamento Ue n. 1689/2024.

L'art. 30 prosegue al comma 2, prevedendo che *“Nell'acquisto o sviluppo delle soluzioni di cui al comma 1 le stazioni appaltanti e gli enti concedenti:*

a) assicurano la disponibilità del codice sorgente, della relativa documentazione, nonché di ogni altro elemento utile a comprenderne le logiche di funzionamento;

b) introducono negli atti di indizione delle gare clausole volte ad assicurare le prestazioni di assistenza e manutenzione necessarie alla correzione degli errori e degli effetti indesiderati derivanti dall'automazione”.

Questo comma fa alcune precisazioni in materia, al fine di assicurare il rispetto del principio di trasparenza. In particolare, impone alle stazioni appaltanti di fornire il codice sorgente (di cui si è accennato sopra), nonché di ogni altro elemento utile a consentire di comprendere come funziona il sistema di intelligenza artificiale (l'algoritmo) utilizzato dalla stazione appaltante.

Questa norma è il frutto dell'evoluzione giurisprudenziale che, nel corso del tempo, ha portato a comprendere che l'interesse non è solo nel risultato fornito dall'algoritmo, quanto nel funzionamento dell'algoritmo stesso e negli *input* che allo stesso vengono forniti. E, così, non ci sarebbe trasparenza e comprensione se il sistema venisse tenuto segreto.

La norma prevede anche che è onere della stazione appaltante introdurre clausole per prevedere alla manutenzione e assistenza dell'attività di automazione per evitare, prevenire o correggere gli errori che dovessero derivarne.

Anche questa previsione è frutto dell'elaborazione giurisprudenziale sviluppatasi in tutti quei casi e quelle procedure dove per un problema tecnico il concorrente non riusciva a presentare offerta nei termini di legge o a caricare l'intero progetto, risultando la sua offerta come tardiva o incompleta per questioni legate al *software* usato dalla stazione appaltante per gestire la piattaforma di ricezione delle offerte.

Prosegue la normativa, prevedendo al comma 3 che *“Le decisioni assunte mediante automazione rispettano i principi di:*

a) conoscibilità e comprensibilità, per cui ogni operatore economico ha diritto a conoscere l'esistenza di processi decisionali automatizzati che lo riguardano e, in tal caso, a ricevere informazioni significative sulla logica utilizzata;

(6) V. in materia <www.primogrado.com/intelligenza-artificiale-e-diritto-amministrativo>.

b) non esclusività della decisione algoritmica, per cui comunque esiste nel processo decisionale un contributo umano capace di controllare, validare ovvero smentire la decisione automatizzata;

c) non discriminazione algoritmica, per cui il titolare mette in atto misure tecniche e organizzative adeguate al fine di impedire effetti discriminatori nei confronti degli operatori economici”.

Questo comma rappresenta, probabilmente, il nucleo centrale della norma perché, oltre ad esprimere i principi che regolano l'intelligenza artificiale, dà anche una descrizione delle possibili tematiche delicate in materia.

In particolare, la lett. a) fa riferimento, ancora una volta, al tema dell'accessibilità sia all'utilizzo del sistema automatizzato, sia alla “logica utilizzata” per far funzionare il sistema stesso. Un aspetto che deve far riflettere, considerando che la portata di questa disposizione dovrebbe valere non solo nell'ambito dei contratti pubblici, ma in ogni procedimento amministrativo.

La lett. b) affronta uno dei temi fondamentali della materia: il rapporto tra la decisione automatizzata e l'uomo che deve controllare il risultato. In particolare, si parla di “non esclusività” della soluzione algoritmica, dimostrando che il semplice fatto che la soluzione deriva da un procedimento automatizzato non sia di per sé garanzia di correttezza o di affidabilità dell'*output* fornito. Spetterà, poi, all'uomo controllare, validare ovvero smentire la decisione automatizzata fornita.

Il problema che si pone al riguardo è legato all'esistenza o meno, all'interno delle pubbliche amministrazioni, di personale capace di gestire i sistemi automatizzati al punto di rendersi conto quando un determinato *output* risulti viziato o errato.

Forse anche per questo il legislatore ha previsto che uno dei principi fondamentali del nuovo Codice è quello della qualificazione delle stazioni appaltanti, tanto che solo le pubbliche amministrazioni che raggiungono certi livelli di qualificazione possono divenire stazioni appaltanti.

La lett. c) replica, sul piano delle decisioni automatizzate, il principio dell'uguaglianza e non discriminazione tutelato anche a livello costituzionale.

In particolare, il legislatore del nuovo Codice richiede che il titolare dell'attività automatizzata si deve fare garante di evitare effetti discriminatori nei confronti degli operatori economici che dovessero derivare proprio dall'uso dell'algoritmo stesso.

Dato che questa norma non prevede sanzioni, non è chiaro quale sia la conseguenza di una mancata attuazione di certe misure preventive in materia discriminatoria, anche se probabilmente potrebbe costituire un nuovo motivo di impugnazione davanti alle giurisdizioni amministrative.

Il comma 4 dell'art. 30 prevede che “*Le stazioni appaltanti e gli enti concedenti adottano ogni misura tecnica e organizzativa atta a garantire che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori, nonché a impedire effetti discriminatori nei confronti di persone fisiche sulla base della nazionalità, dell'origine etnica, delle opinioni politiche, della religione, delle convinzioni personali, dell'appartenenza sindacale, dei caratteri somatici, dello status genetico, dello stato di salute, del genere o dell'orientamento sessuale*”.

Questo comma è volto a tutelare la correttezza dei dati di *input* e che tali inesattezze si riflettano su una lesione dei diritti fondamentali tutelati a livello costituzionale. Difatti, uno dei rischi maggiore nell'inserire dati come *input* di un'intelligenza artificiale è che gli stessi si portino appresso anche informazioni private e non divulgabili relative a privati cittadini.

Anche in questo caso la norma non pone una sanzione, ma è possibile intuire che nel caso di lesione dovuta alla diffusione di dati sensibili ci si muova nella materia della tutela della privacy davanti al giudice ordinario.

Infine, il comma 5 prevede che “*Le pubbliche amministrazioni pubblicano sul sito istituzionale, nella Sezione “amministrazione trasparente”, l'elenco delle soluzioni tecnologiche di cui al comma 1 utilizzate ai fini dello svolgimento della propria attività*». Questo comma si ricollega al tema della trasparenza, imponendo la pubblicazione delle soluzioni tecnologiche scelte sul proprio sito istituzionale.

Come può vedersi, quello che il legislatore può prevedere è una serie di regole e principi generali che dovrebbero guidare la materia. Non è, invece, possibile per il legislatore disciplinare tecnicamente l'uso degli strumenti di intelligenza artificiale. Questo, infatti, rientra nelle competenze degli specialisti tecnici del settore.

Aspetto che crea un problema di coordinamento tra il mondo giuridico e quello dei programmatori o ingegneristico. Infatti, così come il giurista è distante dai tecnicismi del funzionamento degli algoritmi e avrebbe difficoltà a programmare un sistema di intelligenza artificiale, dall'altra parte, i programmatori sono lontani dal mondo del diritto e potrebbero non avere la consapevolezza di quali siano gli *input* corretti da fornire al sistema e quelli che, invece, potrebbero creare problemi in tema di riservatezza del dato personale.

Pertanto, si ritiene necessaria una correlazione e un coordinamento tra l'attività del giurista e quella del tecnico al fine di fornire uno strumento automatizzato e un algoritmo sicuro e realmente efficiente.

In tal senso è auspicabile che anche in sede di approvazione del disegno di legge attualmente al Senato e dei conseguenti decreti attuativi, si avvii una seria interazione tra il mondo giuridico e quello tecnico, attraverso corsi di alfabetizzazione digitale con studio degli aspetti tecnici degli strumenti automatizzati, al fine di creare dei giuristi tecnici e dei tecnici giuristi.

* * *

UNA RIFLESSIONE SISTEMATICA: L'INTELLIGENZA ARTIFICIALE PUÒ SOSTITUIRE UN GIUDICE O TRASFORMARLO DA *HOMO SAPIENS* IN *HOMO NUMERICUS*?

di Vito Tenore (*)

Abstract: Lo studio, alla luce del Regolamento europeo sull'Intelligenza artificiale (Reg. 2024/1689), della prima giurisprudenza e dei tanti studi intervenuti, analizza il problema delle crescenti e sempre più diffuse applicazioni di intelligenze artificiali (AI) in molti campi: medicina, ingegneria, architettura, musica, arte, fotografia, *marketing*, didattica, ricerca, per poi affrontare i vantaggi e gli svantaggi dell'uso dell'AI nel campo giuridico. Redigere una memoria per un avvocato, un rogito per un notaio, un provvedimento amministrativo per un dirigente, una sentenza per un magistrato attraverso intelligenze artificiali, se agevola nello studio e nella ricerca e attenua fazioni e lentezze gestionali, sviscerisce il ruolo centrale del "pensare" dell'essere umano e fa scomparire il buon senso, l'equilibrio, l'etica, il libero arbitrio e, soprattutto, la coscienza di cui è titolare solo l'Uomo, e dunque l'Uomo-Magistrato. L'intelligenza artificiale genera parole senza pensare. La giustizia predittiva, deduttiva o induttiva, urta dunque con il pensare umano. La Giustizia, a differenza di un algoritmo che tende al ragionamento esatto, deve tendere alla verità, coniugando la verità sostanziale con la verità processuale. E questa prerogativa valutativa è solo e soltanto dell'Uomo. Pertanto, nel giudicare in un tribunale, l'intelligenza artificiale non può che avere un ruolo ancillare e servente rispetto alla riflessione e valutazione del caso concreto, che deve restare appannaggio dell'Uomo reale.

The study, in light of the new AI European Regulation (Reg. 2024/1689), the first case law and the several studies overcome in relation to the matter, analyzes the issue concerning the growing and increasing applications of artificial intelligence (AI) in many fields: medicine, engineering, architecture, music, art, photography, marketing, teaching, and research, and then addresses the issue related to advantages and disadvantages of using AI in the legal field. Using artificial intelligence for drafting a memorandum for a lawyer, a deed for a notary, an administrative measure for an executive, a judgment for a Magistrate, while facilitating the study and the research and mitigates factionalism and slow management, debases the crucial role of "thinking" of the human being and displaces common sense, balance, ethics, free will and, above all, the conscience which only Human have, and therefore the Human-Magistrate. AI generates words without thinking. Predictive, deductive or inductive justice thus bumps up against human thinking. Justice, differently from an algorithm that tends to an exact reasoning, must tend to truth, combining substantive truth with procedural truth. And exclusively humans have this evaluative prerogative. Therefore, in the context of a trial before a Court of Law, artificial intelligence can only have an ancillary and servant role with respect to the reflection and evaluation of the concrete case, which must remain a prerogative of humans.

Sommario: 1. *Premessa sul crescente utilizzo dell'intelligenza artificiale (AI) nella vita quotidiana e nel campo giuridico in particolare: vantaggi e rischi.* – 2. *La tendenziale utilità dell'intelligenza artificiale per la società, ma con alcune regole di base. Il Regolamento europeo sull'Intelligenza artificiale (Reg. 2024/1689) entrato in vigore il 2 agosto 2024: la sua finalità preventiva fondata sulla mitigazione del rischio.* – 3. *I rischi sociali, economici e giuridici (civili, penali, lavoristici, processuali) nascenti dal crescente utilizzo dell'intelligenza artificiale, che non è un soggetto di diritto. In attesa di una direttiva europea sui danni da AI.* – 3.1. *Segue: Contenziosi amministrativi e contabili nascenti dall'uso di intelligenza artificiale.* – 4. *Sulla necessità di non abdicare, anche in giudizio, alla primazia dell'intelligenza non artificiale, ma naturale in quanto umana. Dubbi su una sentenza algoritmica.*

1. *Premessa sul crescente utilizzo dell'intelligenza artificiale (AI) nella vita quotidiana e nel campo giuridico in particolare: vantaggi e rischi*

Non è infrequente, dialogando con avvocati, che questi ultimi si lamentino di decisioni irragionevoli, o superficiali, o contraddittorie, o stravaganti e imprevedibili in quanto contrarie a pacifici indirizzi giurisprudenziali assunte da taluni magistrati (1). Non è specularmente infrequente che qualche magistrato critichi la fazione, o la

(*) V. Tenore è Presidente di sezione della Corte dei conti e docente presso la Scuola nazionale dell'amministrazione-Sna. Lo studio, con alcuni adattamenti e aggiornamenti, riproduce il testo del volume V. Tenore, *L'intelligenza artificiale può sostituire un giudice? I rischi della involuzione verso un "cretino digitale"*, Roma, Anicia, 2023.

(1) In effetti talune, in verità rare, decisioni confermano l'antico adagio romanistico *Habent sidera sua lites*, di cui si ha una magistrale interpretazione da parte dell'attore Vincenzo Talarico (nel ruolo del retorico avv. Fabrizio Bartolini) nel film di Steno del 1954 *Un giorno in Pretura*, reperibile in <www.youtube.com/watch?v=wy4t_A-nYCI>.

superficialità, o la temerarietà (a fronte di pacifici indirizzi giurisprudenziali contrari), o la approssimazione di alcuni scritti difensivi o di tesi sviluppate oralmente in udienza da avvocati.

I rischi di questi errori, forzature e stravaganze interpretative o di queste superficialità argomentative o nelle ricerche giurisprudenziali e dottrinali, se sono più tollerabili qualora poste in essere da un avvocato, che deve talvolta, a tutela del cliente, forzare il dato normativo, creare nuove letture, o fingere di ignorare avversi indirizzi giurisprudenziali, non lo sono affatto per il magistrato, le cui sentenze, i cui esiti decisionali devono essere ragionevoli, argomentati, approfonditi e, soprattutto, *prevedibili* secondo il vigente quadro normativo e giurisprudenziale.

Del resto, è la stessa Carta costituzionale, all'art. 3, ad evidenziare, più che in altri articoli, la visione di un diritto oggettivo e certo che deve permeare l'intero ordinamento: la norma impone di trattare in modo uguale situazioni giuridiche uguali, assicurando il medesimo risultato (stesso trattamento) a parità di variabili (medesima situazione).

Se questo è vero, i rischi di imprevedibili o superficiali evenienze decisionali che la realtà giudiziaria talvolta evidenzia, potrebbero sicuramente attenuarsi attraverso lo studio più approfondito (che richiede tempo e fatica) delle carte giudiziarie, della normativa (2), della giurisprudenza (soprattutto della Cassazione che ha una basilare funzione nomofilattica (3), al pari dell'Adunanza plenaria del Consiglio di Stato e delle Sezioni riunite della Corte dei conti) e della dottrina, fermo restando il ragionevole libero arbitrio interpretativo del magistrato nel recepire e sviluppare, nel caso specifico, una propria tesi, anche innovativa, ma pur sempre ancorata al basilare referente normativo e ai prevalenti o addirittura univoci indirizzi giurisprudenziali.

Ma il rischio di stravaganze decisionali o di superficialità argomentative potrebbe essere attenuato seguendo un'altra strada: *l'intelligenza artificiale* (AI) (4), passando, secondo alcuni entusiasti propugnatori di un asserito

(2) Come è noto, il giudice non può discostarsi dalla legge, interpretandola in modo arbitrario, perché violerebbe la legge sull'interpretazione (*ex art. 12 preleggi*) a cui è pienamente assoggettato *ex art. 101 Cost.* L'art. 101 Cost. e l'art. 65 ord. giud. esprimono l'impersonale oggettività del diritto e la funzionalità tecnica della sua applicazione, vietando pre-giudizi e pre-comprensioni.

(3) L'art. 65 ord. giud., nell'indicare le attribuzioni della Corte Suprema di cassazione afferma che questa “*assicura l'esatta osservanza e l'uniforme interpretazione della legge, l'unità del diritto oggettivo nazionale, il rispetto dei limiti delle diverse giurisdizioni*”; l'organo supremo è dunque deputato ad interpretare il diritto, deve assicurare uniformità e unità del diritto oggettivo nazionale.

(4) L'intelligenza artificiale (in sigla IA) è una disciplina che studia se e in che modo si possano realizzare sistemi informatici intelligenti in grado di simulare la capacità e il comportamento del pensiero umano. Secondo la definizione di M. Somalvico, “*L'intelligenza artificiale è una disciplina appartenente all'informatica che studia i fondamenti teorici, le metodologie e le tecniche che consentono la progettazione di sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana*”. In altre parole, l'intelligenza artificiale cerca di mettere in grado i computer di fare il genere di cose che sanno fare le menti umane.

In realtà esistono vari tipi di intelligenza artificiale: quella classica, simbolica spesso definita Gofai, le reti neurali artificiali, la programmazione evolutiva, gli automi cellulari e i sistemi dinamici.

Sul tema dell'intelligenza artificiale in generale e sulle sue ricadute sulla realtà, la politica, la società, la letteratura è assai ampia. Senza pretesa di esaustività, oltre agli scritti citati nelle successive note su profili specifici, si segnalano da ultimo per la loro portata grandangolare: G. Maira, *Intelligence e intelligenza artificiale: cervello, pensiero critico e conoscenza*, prolusione per la Cerimonia di inaugurazione dell'Anno accademico 2024 della Scuola di formazione del Sistema di informazione per la sicurezza della Repubblica, in <<https://www.sicurezza nazionale.gov.it>>; N. Bostrom, *Superintelligenza. Tendenze, pericoli, strategie*, Torino, Bollati Boringhieri, 2023; G. Cassano, G. Scorza (a cura di), *Metaverso. Diritti degli utenti, piattaforme digitali, privacy, diritto d'autore, profili penali, blockchain e NFT*, Pisa, Pacini, 2023; F. Sarzana di Sant'Ippolito, M.G. Pierro, I.O. Epicoco, *Il diritto del Metaverso. NFT, DeFI, GameFI e privacy*, Torino, Giappichelli, 2022; H. Kissinger, E. Schmidt, D. Huttenlocher, *L'era dell'intelligenza artificiale. Il futuro dell'identità umana*, Milano, Mondadori, 2022; F. Faggin, *Irriducibile. La coscienza, la vita, i computer e la nostra natura*, Milano, Mondadori, 2023; D. Cohen, *Homo Numericus. La futura civilizzazione*, Roma, Anicia, 2022; A.F. Uricchio, G. Riccio, U. Ruffolo (a cura di), *Intelligenza artificiale tra etica e diritti. Prime riflessioni a seguito del libro bianco dell'Unione europea*, Bari, Cacucci, 2021 (gli A. promuovono studi integrati, provando a superare le costrittive barriere dei troppo parcellizzati settori scientifici e disciplinari, attraverso una efficace interazione tra le scienze coinvolte); J. Kaplan, *Intelligenza artificiale*, Roma, Luiss University Press, 2021; S. Quintarelli (a cura di), *Intelligenza artificiale. Cos'è davvero, come funziona, che effetti avrà*, Torino, Bollati Boringhieri, 2020; M.A. Boden, *L'intelligenza artificiale*, Bologna, il Mulino, 2019, alla cui ampia bibliografia finale internazionale è sufficiente rinviare; M. Tegmark, *Vita 3.0. Essere umani nell'era dell'intelligenza artificiale*, Milano, Cortina, 2018; M. Buscema, G. Didoné, M. Pandin, *Reti neurali autoriflessive. Teoria, metodi, applicazioni, confronti*, Roma, Armando, 1994; M. Buscema, *Reti neurali artificiali per l'orientamento professionale. Progetto Loop*, Milano, Angeli, 2002.

Interessante è anche la prospettiva di una attenta sociologia dei processi culturali e comunicativi, E. Esposito, *Comunicazione artificiale. Come gli algoritmi producono intelligenza sociale*, Milano, Bocconi University Press, 2022, la quale evidenzia che se le macchine contribuiscono all'intelligenza sociale non è perché hanno imparato a pensare come noi, ma perché hanno imparato a partecipare alla comunicazione. Dobbiamo quindi pensare alle macchine “intelligenti” non in termini di intelligenza artificiale ma in termini di comunicazione artificiale. E per fare questo abbiamo bisogno di un concetto di comunicazione che sappia prendere in considerazione la possibilità che il partner con cui interagiamo non sia un essere umano ma un algoritmo.

Giova infine ricordare il famoso quesito di Alan Turing, padre dell'informatica (a cui è dedicato il film del 2014 *The imitation game*, sul c.d. Codice Enigma): “*Can computers think?*”, dei tardi anni Quaranta del XX secolo, che rimane effettivamente senza risposta, sebbene il test di Turing sia ancora usato per valutare l'output del computer sulla scala dell'intelligenza umana.

“progresso” fondato su *algoritmi* (5), da un suo ruolo ad oggi ancillare rispetto al ragionamento umano, ivi compreso quello giuridico, ad un ruolo preponderante o addirittura assorbente e sostitutivo del magistrato, come rimarcato in diversi saggi (6), convegni (7) ed inaugurazioni di Anni giudiziari.

(5) L'*algoritmo*, sin dalla nozione desunta dai papiri di Ahmes del XVII secolo a.C., consiste in una sequenza finita e ordinata di operazioni elementari e chiare di calcolo che permettono di risolvere, in maniera determinata, un problema. Si tratta di una procedura di calcolo ben definita che consente, attraverso un insieme di operazioni effettuate in un determinato ordine, partendo da un insieme di dati (*input*), di ottenere un risultato atteso (*output*): come si vedrà nel successivo par. 3.1., lo stesso, applicato ad una decisione amministrativa, in via informatica è in grado di valutare e graduare una moltitudine di domande (cfr. Cons. Stato, Sez. VI, n. 2270/2019). La prospettiva, dunque, non è solo quella della semplificazione, ma anche quella della buona amministrazione; alle tecnologie si guarda non solo in vista del miglioramento del processo decisionale, ma anche della qualità della decisione. Secondo la chiara definizione contenuta in *Wikipedia*, l'algoritmo è un concetto fondamentale dell'informatica, anzitutto perché è alla base della nozione teorica di calcolabilità: un problema è calcolabile quando è risolvibile mediante un algoritmo. Inoltre, l'algoritmo è un concetto cardine anche nella fase di programmazione dello sviluppo di un *software*: preso un problema da automatizzare, la programmazione costituisce essenzialmente la traduzione o codifica di un algoritmo per tale problema in programma, scritto in un certo linguaggio, che può essere quindi effettivamente eseguito da un calcolatore rappresentandone la logica di elaborazione.

Il concetto di algoritmo non rappresenta dunque un'acquisizione recente, in quanto è stato infatti enucleato nel contesto della matematica araba dell'Alto Medioevo. Nella sua connotazione pre-tecnologica, esso è identificato in una procedura che conduce alla soluzione di un problema attraverso un percorso formalizzato che si risolve in una sequenza di passaggi pre-costituiti, ordinati ed univoci, tali da poter essere eseguiti e ripetuti senza margini di scelta arbitraria. Sul tema, L. Lombardi Vallauri, *Neuroni, mente, anima, algoritmo: quattro ontologie*, in Id. (a cura di), *Logos dell'essere logos della norma*, Bari, Adriatica, 1999, 586; R. Borruso, C. Tiberi, S. Russo, *L'informatica per il giurista. Dal Bit a internet*, Milano, Giuffrè, 2009, 207; A. Baldi, D. Mula, *Responsabilità civile e intelligenza artificiale*, in G. Taddei Elmi, A. Contaldo (a cura di), *Intelligenza artificiale. Algoritmi giuridici. Ius condendum o "fantadiritto"?*, Pisa, Pacini, 2020, 168.

(6) Sul ruolo dell'intelligenza artificiale nel campo giuridico, limitando le citazioni ai soli studi più rilevanti, v. A. Turing, *Computing Machinery and Intelligence*, in *Mind: new series*, 1950, 443, trad. it. *Meccanismo computazionale e intelligenza*, Roma, Città Nuova, 2018; R. Borruso, *La legge, il giudice, il computer*, in D. Limone (a cura di), *Dalla giuritecnica all'informatica giuridica. Studi dedicati a Vittorio Frosini*, Milano, Giuffrè, 1995, 24; G. Sartor, *Intelligenza artificiale e diritto. Un'introduzione*, Milano, Giuffrè, 1996; G. Terracciano, *L'applicazione in campo giuridico delle reti neurali artificiali. Il programma Giurinet*, in *Tar*, 1998, II, 497; V. Buscema, *"Data base" Sql e reti neurali: un felice connubio*, in *Foro amm.-CdS*, 2003, 2443; A. Santosuosso, C. Boscarato, F. Caroleo, *Robot e diritto: una prima ricostruzione*, in *Nuova giur. civ.*, 2012, II, 497; A. Bertolini, *Robots as products: The case for a realistic analysis of robotic applications and liability rules*, in *Law, innovation and technology*, 2013, 217; G. Sartor, *L'informatica giuridica e le tecnologie dell'informazione. Corso d'informatica giuridica*, Torino, Giappichelli, 2016; E. Palmerini et al., *RoboLaw: Towards a European framework for robotics regulation*, in *Robotics and Autonomous Systems*, 2016, 78; S. Civitarese Matteucci, L. Torchia, *La tecnificazione*, in L. Ferrara, D. Sorace (a cura di), *A 150 anni dall'unificazione amministrativa italiana. Studi*, Firenze, Firenze University Press, 2016; G. Sartor, A. Omicini, *The autonomy of technological systems and responsibilities for their use*, in N. Bhuta et al. (eds.), *Autonomous weapon systems. Law, ethics, policy*, Cambridge, Cambridge University Press, 2016, 39; A. Santosuosso, *Diritto, scienza e nuove tecnologie*, Padova, Cedam, 2016; U. Ruffolo (a cura di), *Intelligenza artificiale e responsabilità*, Milano, Giuffrè, 2017; M. Iaselli, *Robot lawyer: nuovi progressi dell'intelligenza artificiale nel settore legale ma...*, in <www.altalex.com>, 9 novembre 2017; A. Carleo (a cura di), *Calcolabilità giuridica*, Bologna, il Mulino, 2017; B. Fiammella, *Intelligenza artificiale ed etica, tra progresso ed evoluzione*, in <www.altalex.com>, 30 maggio 2018; L. Viola (a cura di), *Giustizia predittiva e interpretazione della legge con modelli matematici* (Atti del convegno tenutosi presso l'Istituto dell'enciclopedia Italiana Treccani), Milano, Centro studi diritto avanzato, 2019; F. Pizzetti, *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, Giappichelli, 2020; U. Ruffolo (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, Giuffrè, 2020; R. Trezza, *Diritto e intelligenza artificiale. Etica, privacy, responsabilità, decisione*, Pisa, Pacini, 2020; E. Calzolaio (a cura di), *La decisione nel prisma dell'intelligenza artificiale*, Padova, Cedam, 2020; A. Santosuosso, *Intelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto*, Milano, Mondadori, 2020; G. Danzi, *Il fattore tempo nella giustizia predittiva e nella certezza del diritto*, in <www.lanuovaproceduracivile.com>, 12 maggio 2021; A. D'Aloia (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Milano, Angeli, 2021; M. Tampieri, *L'intelligenza artificiale e le sue evoluzioni*, Padova, Cedam, 2022; G. Sartor, *L'intelligenza artificiale e il diritto*, Torino, Giappichelli, 2022; E. Bassoli, *Algoritmica giuridica. Intelligenza artificiale e diritto*, s.l., Amon, 2022; R. Giordano et al. (a cura di), *Il diritto nell'era digitale*, Milano, Giuffrè, 2022; L. Di Donna, *Intelligenza artificiale e rimedi risarcitori*, Padova, Cedam, 2022; G. Terracciano, *I.I.A. (intelligenza artificiale amministrativa) e sindacato giurisprudenziale*, in *www.amministrativamente.it*, fasc.2, 2022 D. Buzzelli, M. Palazzo (a cura di), *Intelligenza artificiale e diritti della persona*, Pisa, Pacini, 2022; G.M. Riccio, G. Ziccardi, G. Scorza (a cura di), *Intelligenza artificiale: profili giuridici*, Padova, Cleup, 2022; A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, Bologna, il Mulino, 2022; N. Abriani, R. Costi (a cura di), *Diritto societario, digitalizzazione e intelligenza artificiale*, Milano, Giuffrè, 2023; L. Viola, *La giustizia predittiva del lavoro*, in <www.lavorodirittieuropa.it>, 2023; G. Pesce, *L'intelligenza artificiale alla prova del diritto europeo: verso il diritto della paura?*, in <www.amministrativamente.it>, 2024.

(7) Tra i principali eventi organizzati su tema, oltre a decine di giornate di studio organizzate in contesti universitari, va segnalato il convegno "Intelligenza artificiale amministrativa (IAA). Principi e regole per l'esercizio delle funzioni pubbliche" organizzato dall'Inps e l'Università degli Studi di Roma "Foro Italico" il 4 e il 5 luglio 2024, a Roma nella sede di Palazzo Wedekind, e quello curato da Upel, 7^a Rassegna di Diritto pubblico dell'economia-Intelligenza artificiale e Appalti pubblici tra capacità predittiva e discrezionalità amministrativa, tenutosi a Varese il 18 e 19 aprile 2024. Degno di nota è poi il convegno a cura dalla giustizia amministrativa, "Intelligenza artificiale, diritti, giustizia e pubblica amministrazione", tenutosi a Palazzo Spada, Consiglio di Stato il 18 maggio 2023. Va poi menzionato l'evento curato dall'Avvocatura dello Stato

In particolare, sia in fase decisoria, sia in fase conciliativa (8) del ragionamento giuridico, l'intelligenza artificiale tenta di simulare gli esiti per una vertenza, avendo a disposizione un numero molto alto di precedenti oltre che di tutti i referenti normativi: più alta è la base statistica, maggiore è l'affidabilità dello strumento. Si tratta della *giustizia predittiva induttiva*, che utilizza un metodo associazionistico puro: selezionati i metadati disponibili, esita le soluzioni più frequentate nelle aule processuali e fornisce un utile argomento esperienziale, avvicinando però il nostro sistema italiano alla tecnica dei precedenti, tipica dei Paesi di *common law*, e tale metodica sembra essere stata ripresa anche dal novello art. 425, c. 3, c.p.p. introdotto da riforma Cartabia sulla "*ragionevole previsione di condanna*" (9). Oppure, l'AI può tentare di simulare il ragionamento fondato sul sillogismo giudiziario (o dell'interpretazione delle norme) descritto dall'art. 12 delle disp. prel. c.c. (10): argomenti semantici, sintattici interni alla norma, argomenti analogici o intenzionali del legislatore. Quest'ultima è chiamata *giustizia*

il 30 marzo 2023, sul tema "Algoritmi ed intelligenza Artificiale: il dialogo necessario tra il diritto e le nuove tecnologie". Rilevante anche il corso curato per magistrati dalla Scuola superiore della magistratura a Napoli il 26-28 settembre 2022, sul tema "Intelligenza artificiale e diritto". Si segnalano anche i diversi convegni organizzati dall'Ordine degli avvocati. Ad esempio, l'Ordine di Roma ne ha curati diversi: il 1° giugno 2023 su "Intelligenza artificiale e transizione digitale: l'obbligatoria evoluzione del ruolo del professionista"; il 4 ottobre 2022 su "L'intelligenza artificiale nei sistemi giudiziari. Criticità e opportunità - Il ruolo degli avvocati"; il 1° marzo 2022 su "Intelligenza artificiale e diritto". Anche l'Ordine degli avvocati di Milano ha organizzato il 22 giugno 2023 il convegno "Intelligenza artificiale e privacy per lo studio legale". V. anche il convegno tenutosi presso l'Istituto dell'enciclopedia italiana Treccani, i cui atti sono in L. Viola (a cura di), *Giustizia predittiva e interpretazione della legge con modelli matematici*, cit.

(8) Alternative al contenzioso sono le c.d. ADR, *Alternative dispute resolution*, o Marc, *Metodi alternativi di risoluzione delle controversie*. Una maggiore prevedibilità (*rectius*: predittività) della sentenza può agevolare sistemi di risoluzione alternative alle controversie (c.d. ADR) come mediazione ex d.lgs. n. 28/2010, arbitrato ex art. 806 c.p.c. ss., negoziazione di cui al d.l. n. 132/2014, proposta conciliativa ex art. 185-bis c.p.c., ma anche transazioni: tanto più la sentenza è prevedibile, tanto più le parti saranno indotte a trovare soluzioni alternative che prendano le mosse dalla possibile sentenza come "base per l'accordo o decisione".

(9) Osserva G. Riccio, *Ragionando su intelligenza artificiale e processo penale*, in Arch. pen., 2019, 1, è «usanza comune far riferimento alla formula "giustizia predittiva" per riferirsi alla possibilità di prevedere l'esito di un giudizio tramite calcoli e algoritmi e, quindi, per mezzo degli strumenti dell'intelligenza artificiale. È ovvio. Non si tratta di formule magiche, ma del modo di prevedere la probabile decisione di uno specifico caso, attraverso l'ausilio di algoritmi, e risolvere così la vecchia questione del "precedente", che assume rilievo decisivo nei Paesi di common law e nella giurisprudenza per casi, dove ha valore vincolante, non nei sistemi di civil law, dove riveste, invece, un rilievo scarsamente persuasivo e in rari casi anche convincente». Rimarca però M. Ferrari, *Profili giuridici della predizione algoritmica*, Napoli, Guida, 2022, 86, che "la giurisprudenza [...] non può fondare predizioni, né tanto meno certezze, sulla sola osservazione del passato".

Sul nuovo art. 425, c. 3, c.p.p. (dopo Cartabia) che menziona la "*ragionevole previsione di condanna*", v. C. Intrieri, L. Viola, *Ragionevole previsione di condanna e giustizia predittiva: una modesta proposta per la riforma dell'art. 425 c.p.p.*, in <www.giustiziainsieme.it>, 1° febbraio 2022.

(10) In base all'art. 12 delle preleggi, che è una sorta di "*algoritmo scritto*" (seguendone la stringente logica consequenziale), per interpretare una legge bisogna procedere come segue: a) attribuire il significato letterale, che è la c.d. interpretazione letterale; b) comporre l'interpretazione letterale con l'intenzione del legislatore, che è la c.d. interpretazione per *ratio*; c) in assenza di precisa disposizione letterale, nel senso di assenza di norma oppure di presenza di norma, ma con significato equivoco, si può utilizzare la c.d. *analogia legis*, ovvero cercare nel medesimo complesso di leggi (codice civile, per quello che qui rileva) una situazione giuridica tipizzata analoga, con stessa ratio (c.d. *eadem ratio*); d) se il caso è ancora dubbio, si può procedere ad utilizzare i principi generali dell'ordinamento, che è la c.d. *analogia iuris*.

L'algoritmo potrebbe essere impostato con una analoga sequenza predeterminata di operazioni per giungere ad un risultato ragionevole e affidabile.

Come ben ricorda L. Viola, *La giustizia predittiva del lavoro*, cit., 4, il primo caso di incidenza degli algoritmi sulla giustizia si è avuto con Compas. La questione è stata affrontata dalla Suprema Corte dello Stato del Wisconsin, negli Usa, caso *Stato contro Loomis*, riguardante la legittimità dell'applicazione di un *software*/algoritmo denominato Compas (algoritmo proprietario) per la determinazione della recidività di un soggetto che si era macchiato di un determinato crimine ai fini dell'applicazione della pena.

predittiva deduttiva e tocca il patrimonio di abilità proprie dell'ermeneutica giuridica (11). Alcune sperimentazioni di giustizia predittiva sono state sviluppate in uffici giudiziari nazionali e stranieri (12).

Sin dall'antichità l'uomo ha sentito il bisogno di prevedere il futuro, per essere preparato: nelle società antiche, la capacità di prevedere il futuro era spesso associata a figure religiose o spirituali, come i profeti, i sacerdoti o gli sciamani. Ogni epoca ha avuto il suo "ponte" tra il presente e il futuro per rispondere sempre alla stessa esigenza di controllo e di gestione dell'incertezza. Oggi questa funzione potrebbe essere assolta dalla c.d. giustizia

(11) Sulla giustizia predittiva induttiva e deduttiva e sull'utilizzabilità dell'AI anche in sede Adr, v. L. Viola, *Giustizia predittiva: è preferibile un modello deduttivo*, in <www.altalex.com>, 10 marzo 2020, che reputa, in modo argomentato, più attendibile e coerente con il nostro sistema la giustizia predittiva induttiva; Id., voce *Giustizia predittiva*, in <www.trecani.it>; F.G. Capitani, *Chi vince e chi perde con l'intelligenza artificiale nel diritto*, in <www.dirittoegiustizia.it>, 4 aprile 2023; Id., *Intelligenza artificiale e ADR (e mediazione ex d.lgs. n. 28/2010): un chiarimento di metodo*, ivi, 10 luglio 2023, il quale ricorda che sono già stati prodotti software in grado, nella migliore ipotesi, di fornire un responso o un parere a un quesito o sugli esiti di una controversia; in parte utilizzando criteri di tipo statistico/esperienziale/associazionistico, in misura di gran lunga minore utilizzando logiche affini allo speculare giuridico astratto, la cui raffinatezza trova difficoltà superiori. L'A. rimarca inoltre come "La via dell'AI alla ADR e alla mediazione strizza l'occhio all'economia e, nonostante sia ritenuta in primis il luogo delle convergenze psicologiche fra le parti, si presta in realtà molto di più di altre branche del diritto all'utilizzo di metodi logico matematici ed economici, in genere disprezzati da un'ideologia onnivora del ragionamento giuridico".

L. Viola, *La giustizia predittiva del lavoro*, cit., 10 ss., richiama i risultati di uno studio empirico condotto di recente dal Centro studi diritto avanzato su 12 casi esaminati relativi a ordinanze di rimessione alle Sezioni unite della Cassazione: utilizzando il modello deduttivo, 8 sentenze sono state predette correttamente (66,6%); utilizzando il modello induttivo, 4 sentenze sono state previste correttamente (33,3%).

Sul tema, v. anche M.A. Catarozzo, *Giustizia predittiva e Intelligenza Artificiale: quel bisogno antico di prevedere il futuro*, in <www.dirittoegiustizia.it>, 26 maggio 2023, in quale ben sottolinea come "fin dall'antichità l'uomo ha sentito il bisogno di prevedere il futuro, per essere preparato e per dare una risposta all'ansia legata all'incertezza. Il desiderio di prevedere il futuro è radicato nella storia umana. L'essere umano ha cercato così di comprendere e prevedere eventi del futuro per potersi preparare di fronte a possibili pericoli. Gli agricoltori dipendevano dalla previsione delle stagioni e del clima per sapere quando piantare e raccogliere i loro raccolti; i generali cercavano segni divini per decidere se e quando attaccare in battaglia. Prevedere un evento, fornisce la sensazione (c.d. bisogno di controllo) di poterlo controllare e di poterlo modificare. Nell'antica Roma, i consoli cercavano il consiglio degli aruspici, sacerdoti che leggevano i presagi nel fegato degli animali sacrificati, per prevedere il risultato di una battaglia o di un evento politico"; P. Moro, *Intelligenza artificiale e professioni legali. La questione del metodo*, in *Journal of Ethics and Legal Technologies*, 2019, 24; C. Castelli, D. Piana, *Giustizia predittiva. La qualità della giustizia in due tempi*, in *Questione giust.*, 2018, 4, 153; D. Dalfino, *Creatività e creazionismo, prevedibilità e predittività*, in *Foro it.*, 2018, V, 385; L. Viola, *Interpretazione della legge con modelli matematici*, Milano, Centro studi diritto avanzato, 2018, 167; A.R. Gaglioti, *Un modello matematico per l'analisi quantitativa del diritto*, Reggio Calabria, Città del sole, 2011. Specificamente sull'utilizzo ai fini Adr, v. C.E. Bruno, *Mediazione e prevedibilità della sentenza*, in <www.lanuovaproceduracivile.com>, 15 ottobre 2020.

Dei parametri della giustizia predittiva si fa già oggi frequente uso in tema di responsabilità dell'avvocato, dove è dirimente individuare cosa sarebbe accaduto, in termini di decisione giudiziale, laddove l'avvocato avesse posto in essere l'azione processuale in concreto non espletata: sul punto, G. Spina, *Responsabilità dell'avvocato e giustizia predittiva*, ivi, 28 marzo 2018; P. Calamandrei, *Limiti di responsabilità del legale negligente*, in *Riv. dir. proc.*, 1931, II, 260, nell'affrontare il tema della responsabilità dell'avvocato e nella necessità di calcolare la probabilità di accoglimento di un appello, proponeva di quantificare l'eventuale danno da mancato appello moltiplicando il *petitum* per la percentuale di accoglimento degli appelli in generale. La giurisprudenza tende ad orientare tale accertamento con il sistema del c.d. più probabile che non: cfr. Trib. Roma, 16 gennaio 2019 afferma che "in tema di responsabilità professionale dell'avvocato per omesso svolgimento di un'attività dovuta da cui sarebbe potuto derivare un vantaggio patrimoniale per il cliente in violazione dell'obbligo qualificato di cui all'art 1176 c 2° cc (nel caso di specie la mancata costituzione) occorre verificare non solo la eventuale colpa del difensore ma altresì il nesso di causalità tra la condotta colposa e l'evento di danno secondo un giudizio prognostico sull'esito della eventuale attività omessa secondo il principio del più probabile che non".

Alcune aziende private hanno diffuso software di giustizia predittiva: modello induttivo (OneLegale); modello deduttivo (GiuriMatrix).

(12) Ad esempio, sul sito della Corte d'appello di Venezia è stata attivata una nuova sezione denominata "Giurisprudenza predittiva"; un progetto simile, denominato "Giustizia predittiva", è stato attivato anche dalla Corte d'appello di Brescia; analogo è il progetto "Prevedibilità delle decisioni" della Corte d'appello di Bari. La Corte di cassazione e la Scuola universitaria superiore Iuss di Pavia hanno sottoscritto il 29 settembre 2021 un accordo quadro al fine di attivare una collaborazione strategica, tra la stessa Scuola e il Centro elettronico di documentazione (Ced) della Corte, per lo sviluppo di ricerca avanzata nel settore degli strumenti tecnici per la raccolta e l'organizzazione del materiale giuridico digitale. Sul tema, v. C. Morelli, *Giustizia predittiva: il progetto (concreto) della Corte d'appello di Brescia*, in <www.altalex.com>, 8 aprile 2019. Con riguardo alle Corti europee, è in corso uno studio di giustizia predittiva sulle sentenze della Corte europea dei diritti dell'uomo: il metodo è il primo che cerca di prevedere le decisioni di un tribunale internazionale analizzando automaticamente il testo dei documenti di causa utilizzando un algoritmo ad apprendimento automatico.

In Francia è stata lanciata una piattaforma che mira proprio a prevedere l'esito giudiziale, tramite un calcolo delle probabilità della definizione di una causa, l'ammontare dei risarcimenti ottenuti in contenziosi simili, e identifica gli argomenti su cui vale la pena di insistere. Ovviamente sulla base di informazioni inserite dall'utente e passando in rassegna milioni di documenti, leggi, norme e sentenze (C. Morelli, *op. cit.*). Negli Usa vi sono già due piattaforme in materia: Premonition (che consente di conoscere le decisioni, i precedenti, le citazioni di ogni singolo giudice, quali argomenti o linguaggio il giudice trova più persuasivo, i suoi scritti e le sue opinioni, gli articoli che gli sono dedicati) e Ravel Law.

predittiva affidata ad AI, ai cui algoritmi si potranno fornire poderose masse di informazioni giurisprudenziali, dottrinali, normative (13).

Del resto, come ben rimarcato in dottrina (14), il diritto può essere costruito come una scienza, che trova la sua principale ragione giustificativa nella misura in cui è garanzia di certezza: il diritto nasce per attribuire certezza alle relazioni umane, tramite una complessa attribuzione di diritti e doveri. La giustizia predittiva consentirebbe di prevedere la probabile sentenza, relativa ad uno specifico caso, attraverso l'ausilio di algoritmi.

Oltre alla maggior prevedibilità degli esiti di una lite, con conseguente verosimile deflazione del contenzioso (le liti temerarie si ridurrebbero, ovviamente se gli avvocati spiegassero lealmente ai clienti i rischi del contenzioso "audace") e maggior ricorso a strumenti conciliativi-Adr e a transazioni, il ricorso all'intelligenza artificiale avrebbe inoltre altri pregi, a vantaggio, ad esempio, della riduzione dei criticati tempi lunghi della giustizia: dispensare il magistrato da lunghe ricerche o analisi dei dati fattuali in giudizi voluminosi, o agevolarlo nella stesura in contenziosi seriali e routinari connotati da modesti margini valutativi o di studio, aumenterebbe la produttività e la rapidità. In sintesi, accelererebbe i tempi della giustizia a vantaggio della collettività nelle cause più ripetitive e banali, ovvero meno "cerebrali". E tale conclusione vale anche per l'attività investigativa delle procure e della Procura generale, che potrebbero trarre gran giovamento, sul piano della contrazione dei tempi investigativi e della loro completezza e doverosa profondità, da una gestione intelligente di masse rilevanti di dati giudiziari in moltissime indagini, tema esplorato in recenti convegni (15).

Un terzo vantaggio della giustizia algoritmica è dato dalla sua impermeabilità agli "umori" e ai preconcetti (frutto del proprio portato esistenziale) che influenzano e condizionano i giudizi umani, ivi compresi quelli dei magistrati. Alcuni studi (16) hanno dimostrato, con dati statistici: a) che i giudici comminano pene più severe il lunedì successivo alla sconfitta della propria squadra di calcio; b) che i giudici sono più clementi il giorno del compleanno dell'imputato (non ci sono studi sulla applicabilità del dato al giorno del compleanno del giudice); c) che con l'avvicinarsi dell'orario di pranzo i giudici sono più respingenti su istanze di trattamenti carcerari di *favor* (17); d) che il clima afoso influenza le decisioni, con conseguenti condanne più severe nei giorni di grandi calore; e) che nei concorsi i commissari nelle giornate di tempo bello sono più sensibili alle qualità non accademiche dei candidati. Tutte queste evenienze (alle quali si possono aggiungere diversi altri fattori "umani") (18) non intaccano

(13) Sulla scorta di dati normativi, precedenti giurisprudenziali, dottrina e interpretazioni, e di esiti dei casi legali l'avvocato potrà decidere, basandosi sulle previsioni dell'esito della causa, se procedere con l'azione legale o tentare la strada extragiudiziale o optare per una transazione; potrà valutare quale linea difensiva sarà più appropriata, in funzione delle probabilità di esito positivo; potrà dare risposte preventive alle richieste della clientela e consigliare di conseguenza come comportarsi. Nel contempo, il magistrato, quando rivestirà i panni della pubblica accusa, avrà nuovi strumenti di indagine, di valutazione e di previsione, alla stregua della difesa; da giudice, avrà nuovi strumenti di valutazione e un nuovo alleato nella gestione delle informazioni e dei dati su cui basare la decisione.

Ricorda L. Viola, voce *Giustizia predittiva*, cit.: "*Leibniz, già nel 1666, evidenziava la possibilità di utilizzare modelli di giustizia predittiva, in particolare utilizzando modelli matematici; secondo costui (Leibniz, G.W., Dissertatio de Arte combinatoria, 1666), difatti, davanti ad una disputa, un giorno, non sarebbe stato necessario un processo, ma si sarebbe potuto direttamente procedere ad un calcolo. Egli immaginava, cioè, una calcolabilità delle controversie tramite veri modelli matematici; ciò avrebbe determinato una maggiore prevedibilità. Per Weber (Weber, M., Die Wirtschaft und die gesellschaftlichen Ordnungen, Tübingen, 2001, 40), l'economia moderna può crescere tramite contratti, ma questi esigono un funzionamento del diritto calcolabile secondo regole razionali: la razionalità formale passa dalla calcolabilità completa dell'ordinamento giuridico. Il diritto è basato su regole scritte e la certezza del diritto altro non è se non la prevedibilità dell'esito giudiziale; in fondo, sarebbe pur sempre un diritto razionale formale (Weber, M., Economia e società, II, Milano, 1974, 17)*".

(14) L. Viola, *Giustizia predittiva: è preferibile un modello deduttivo*, cit.

(15) V. l'interessante convegno tenutosi il 12 luglio 2023 a Milano presso la Fondazione Giangiacomo Feltrinelli su "L'intelligenza artificiale nello svolgimento delle funzioni requirenti. Un dibattito in chiave prospettica sull'impiego dell'IA al servizio della Giustizia".

(16) D. Kahneman, O. Sibony, C.R. Sunstein, *Noise, A flaw in human judgment*, London, William Collins, 2022.

(17) Lo dimostra uno studio di S. Danziger, J. Levav, L. Avnaim-Pesso, *Extraneous factors in judicial decisions*, in <www.pnas.org>, 25 febbraio 2011, che prende in esame una serie di sentenze per la concessione della libertà condizionata, emesse da giudici israeliani di comprovata esperienza. Documenta lo studio che le due interruzioni giornaliere dell'attività lavorativa, per uno *snack* e per la pausa pranzo, di fatto segmentano la giornata in tre distinte "sessioni". Infine, rileva il saggio che la percentuale di pronunciamenti favorevoli alla concessione della libertà condizionata diminuisce gradualmente, dal 65% fin quasi allo zero, in ognuna di queste sessioni per poi risalire bruscamente allo stesso 65% dopo una pausa lavorativa. Viene dunque constatato che le decisioni di un giudice possono essere condizionate da variabili esterne che non devono invece avere alcuna rilevanza sul giudizio espresso.

(18) Si pensi a dolorose esperienze private, analoghe a quelle sottopostegli in giudizio, fatte dal magistrato stesso (una sofferta ed "espoliativa" separazione o un divorzio; una dura lite condominiale su questione analoga; una vicenda di intollerabili e reiterate immissioni acustiche similari patite, ecc.), a simpatie o antipatie epidermiche o legate a fatti extralavorativi verso una delle parti (non tali da giustificare l'astensione o recusazione, ma influenzanti a livello subliminale), al proprio portato esistenziale (una vecchia "fiamma" adolescenziale o un avversario sportivo è parte in giudizio) o di salute (una analoga malattia di cui è vittima una parte affligge uno stretto parente del giudice), alla propria appartenenza ideologica, correntizia o lobbistica, ecc. Tali fattori potrebbero alterare un lineare e indipendente giudizio.

invece l'imparziale giudizio di una AI, insensibile a variazioni climatiche e appetito, ad antipatie o simpatie, a pulsioni, ad ideologie e a pregresse esperienze umane.

In realtà, il "delegare" ad una macchina il parziale lavoro di redazione di atti giudiziari, compresa una sentenza, e dunque di "decidere" (ma lo stesso vale per attività consultive e difensive di avvocati, o per istruttorie di procure, o per rogiti notarili), è un fenomeno da decenni diffuso ove si dia una definizione ampia di intelligenza artificiale atta a ricomprendere, oltre ai formulari telematici sempre più diffusi (non solo tra avvocati, ma anche tra magistrati e notai) (19), soprattutto i lavori basici di onerosa ricerca di precedenti giurisprudenziali o di scritti dottrinali, delegati non più a collaboratori o a letture personali di riviste e repertori cartacei (20), ma a motori di ricerca (aperti a tutti o contenuti in banche dati ad accesso riservato) più o meno complessi o completi, in cui inserire dati normativi, parole chiave, precedenti giurisprudenziali. Il testo di tali precedenti giurisprudenziali (e dottrinali), reperiti, ad oggi, dal combinato disposto di uomo e motore di ricerca (a sua volta realizzato, riempito e aggiornato da uomini), vengono poi riutilizzati, attraverso agevoli "copia e incolla", dal magistrato (o dall'avvocato) nel redigere la propria sentenza (o nella propria comparsa). Il testo così redatto (*rectius* "assemblato"), pur perdendo di fluidità ed originalità linguistica ed espositiva e pur reiterando concetti e "frasi-tipo" coniati in un precedente *leading-case*, risulta più rapidamente realizzato, sicuramente con minore sforzo cerebrale, concetto sul quale torneremo, in chiave problematica, nella parte finale di questo scritto, evidenziando i rischi di una eccessiva delega all'AI della riflessione umana in generale, e soprattutto nella complessa redazione di una sentenza.

Ma già da tempo si è passati in molti campi da un utilizzo del computer come mero archivio, ad un uso dello stesso come sostituto del funzionario, del professionista, e, dunque, di processi mentali propri dell'Uomo.

Dunque, il salto da tale attuale limitato ruolo servente e ancillare dell'intelligenza artificiale quale mero motore di ricerca di una banca dati, ad una più ampia delega da parte dell'uomo-magistrato di funzioni di studio, elaborative, motivazionali e decisionali di una sentenza ad una AI (e lo stesso vale per una analoga delega data da un avvocato ad una macchina per rendere un parere o redigere un atto difensivo, o da un p.m. per formulare una automatizzata richiesta di rinvio a giudizio o depositare conclusioni, o da un notaio per redigere atti), sebbene da molti auspicato e da molti studiato, pone rilevanti problemi etici, sociali, filosofici, di evoluzione o di scomparsa di lavori e professioni. Ma tale salto pone anche nuovi e rilevanti problemi giuridici di ascrizione delle responsabilità in caso di errori commessi dall'AI, tematica già esplorata a livello dottrinale e persino giurisprudenziale in taluni Paesi, come si vedrà nel paragrafo 3.

Il problema si è già posto in giurisprudenza in relazione all'uso di algoritmi per l'adozione di atti amministrativi in violazione di regole basiche sul procedimento amministrativo, quali la motivazione e la partecipazione procedimentale e la trasparenza (21).

Se poi questa delega di funzioni cerebrali umane fosse totale, con conseguente scomparsa di un giudice, di un avvocato, di un notaio o di un dirigente pubblico (preposto a gare, concorsi, gestione del personale, istruttorie procedimentali, ecc.) "persone fisiche", sostituiti da una "*res intelligente*", ovvero da un "*robot togato*" capace di autoprogrammarsi e aggiornarsi, la situazione, non certo avveniristica o fantascientifica, renderebbe il tema ancor più complesso.

Ad oggi l'intelligenza artificiale sta avendo vasta applicazione attraverso il GPT-3 creato da Open AI e la sua applicazione Chat-GPT (22), che è in grado di generare testi, nuovi contenuti, come immagini, video e testi, senza

(19) Tali formulari sono l'omologo del "precedente" o dello "stampone" utilizzato da sempre dai dirigenti pubblici (ma anche privati) per redigere atti.

Sul tema dei contratti redatti con utilizzo di AI, v. F. Parisi, *Il contratto concluso mediante computer*, Padova, Cedam, 1987; E. Tosi, *I contratti di informatica. Tipi contrattuali, formazione e responsabilità*, Milano, Pirola, 1993; Id., *Il contratto virtuale. Procedimenti formativi e forme negoziali tra tipicità e atipicità*, Milano, Giuffrè, 2005; Id., *Diritto privato dell'informatica e di internet*, Milano, Giuffrè, 2006; R. Clarizia, *I contratti informatici*, in P. Rescigno (diretto da), *Trattato dei contratti*, Torino, Utet, 2007.

(20) Chi non ricorda con nostalgia le lunghe giovanili ricerche cartacee in biblioteca per tesi di laurea, per redazione di articoli e note, e poi, più maturi, per cercare, da neo-magistrati o da avvocati, precedenti giurisprudenziali nei poderosi repertori cartacei, elegantemente rilegati, di blasonate riviste, oggi quasi scomparse o trasformate in riviste telematiche con proprio motore di ricerca. Chi non ricorda con ammirazione e rispetto anziani colleghi depositari di poderosa cultura giuridica e dalla prodigiosa memoria, in grado di fornire, a braccio, ausili "cerebrali reali" in camere di consiglio su casi analoghi o precedenti giurisprudenziali recenti o risalenti.

(21) Sul punto, v. il successivo par. 3.1.

(22) Trattasi di un *chatbot* basato su intelligenza artificiale e apprendimento automatico che gestisce conversazioni con utenti umani. La sigla GPT sta per *Generative Pretrained Transformer*, la tecnologia di apprendimento usata dalla macchina. La causa californiana non è l'unica intentata contro Open AI ed altre start-up operanti nel campo dell'AI: lo ricorda V. Della Sala, *Dalle stelle alla class action: Chat GPT accusata di furto*, in *Il Fatto quotidiano*, 2 luglio 2023, 17.

In Italia va segnalato che il Garante della privacy, con provvedimento 30 marzo 2023 n.112 (seguito da provv.to 11 aprile 2023), prendendo atto che "*non viene fornita alcuna informativa agli utenti, né agli interessati i cui dati sono stati raccolti da OpenAI, L.L.C. e trattati tramite il servizio di ChatGPT; [...] l'assenza di idonea base giuridica in relazione alla raccolta dei dati personali e al loro trattamento per scopo di addestramento degli algoritmi sottesi al funzionamento di ChatGPT;*

l'intervento umano, addestrandosi su un ampio set di esempi e imparando a creare nuovi esempi simili, diversi da quelli visti in precedenza e ponendo problemi di diritto d'autore sui testi elaborati da una macchina, che resta pur sempre una *res*, come tale inidonea a rivendicare propri personali diritti, su opere definibili "orfane" di un proprietario. Per gli esiti valutativi di un esame medico, l'analogo problema è risolto attraverso il recepimento delle risultanze del referto per immagine e per dati numerici da parte del medico che si è valso dello strumento (Tac, risonanza magnetica, radiografia, esiti di esami del sangue forniti da macchine con reagenti, ecc.).

Nei confronti di Open AI, creatrice di Chat-GPT, è stata però di recente intentata in un tribunale federale californiano, a luglio 2023, dallo studio californiano Clarkson, una *class action* per asserita violazione del diritto d'autore e della *privacy* di milioni di cittadini, in quanto il sistema di intelligenza artificiale "generativa" fa allenare i suoi algoritmi con masse di dati e informazioni acquisibili in fonti aperte *online*, ovvero quelle consultabili da chiunque. La rielaborazione di questi dati da parte dell'AI (con capacità superiori a quelle dell'uomo) per fornire testi, immagini, opere di ingegno, comporta, secondo i legali promotori dell'azione giudiziaria, una violazione del diritto d'autore e della *privacy* degli originari autori dei dati riutilizzati. In assenza di solidi riferimenti normativi in materia, sarà la giurisprudenza a fissare limiti all'acquisizione e rielaborazione di dati aperti. Altra azione è stata sporta il 27 dicembre 2023 dal New York Times davanti alla Corte di New York nei confronti di diverse società appartenenti al gruppo OpenAI e Microsoft per violazione dei diritti d'autore sulle proprie pubblicazioni giornalistiche, oltre che per concorrenza sleale e per diluizione del marchio.

Anche nell'ambito dell'insegnamento universitario, in altri Paesi sono state introdotte forme di AI al servizio degli studenti: l'Università di Harvard ha messo a disposizione degli studenti CS50, una intelligenza artificiale analoga a Chat-GPT, disponibile 24 ore su 24 ogni giorno, come assistente, guadagnando forse in disponibilità di dati offerti e in tempi dedicati ai discenti, ma perdendo ovviamente quella "empatia personale" che rappresenta il profilo più arricchente del rapporto tra "maestro e allievo".

Ma moltissimi sono i campi d'azione, attuali e futuri, per l'intelligenza artificiale, oggetto pertanto di diversi corsi di laurea, come ben schematizzato nel *Programma strategico intelligenza artificiale 2022-2024*, redatto dal Ministero dell'università e della ricerca, dal Ministero dello sviluppo economico e dal Ministro per l'innovazione tecnologica e la transizione digitale, del 24 novembre 2021 (23).

Questa crescente presenza di macchine intelligenti è stata ormai univocamente accettata per lavori faticosi, routinari, alienanti, o pericolosi: nelle catene di montaggio industriali, nella gestione di magazzini, in *chat* con risposte a quesiti ricorrenti posti in via telematica a imprese di servizi (banche, assicurazioni, gestori telefonici o di TV, fornitori di servizi pubblici o privati, ecc.). È inoltre accettata quale (pericoloso) sistema di ricerca in Google o quale strumento di gioco (ad es. per partite di scacchi) (24).

Parimenti accettata è la progressiva introduzione di intelligenze artificiali e macchine evolute nel comparto sanitario, in quanto più affidabili (per maggior precisione nei movimenti, per più rilevanti bagagli esperienziali pregressi e per capacità più prolungata di lavoro) di alcune diagnosi o di alcuni gesti umani nell'interventistica. Si è dunque passati "dal bisturi al joystick" e un chirurgo deve sapere affrontare una problematica chirurgica con tutte le metodiche a disposizione: chirurgia tradizionale, chirurgia laparoscopica, chirurgia robotica (25).

[...] che il trattamento di dati personali degli interessati risulta inesatto in quanto le informazioni fornite da ChatGPT non sempre corrispondono al dato reale", ha disposto, ai sensi dell'art. 58, par. 2, lett. f), del Regolamento sulla *privacy*, in via d'urgenza, nei confronti di Open AI I.I.C., società statunitense sviluppatrice e gestrice di ChatGPT, in qualità di titolare del trattamento dei dati personali effettuato attraverso tale applicazione, la misura della limitazione provvisoria, del trattamento dei dati personali degli interessati stabiliti nel territorio italiano, richiedendo ad Open AI con provv.to 11 aprile 2023: a) di prevedere il consenso dell'interessato quale presupposto per l'utilizzazione di dati personali, b) la possibilità di rettificare dati inesatti, c) garanzie a tutela dell'accesso di minorenni senza il consenso dei genitori.

Open AI si è conformato a tali richieste ed il Garante ha nuovamente autorizzato l'accesso agli utenti italiani a ChatGPT.

(23) I settori di operatività delle AI sono individuati nel *Programma strategico* alle pp. 17-18: Industria e manifatturiero; Sistema educativo; Agroalimentare; Cultura e turismo; Salute e benessere; Ambiente, infrastrutture e reti; Banche, finanza e assicurazioni; Pubblica amministrazione; Città, aree e comunità intelligenti; Tecnologie dell'informazione; Sicurezza nazionale. Il programma è reperibile in <<https://innovazione.gov.it>>.

E' ormai diffusissimo studiare l'intelligenza artificiale nelle università italiane: difatti sono ben 45 i corsi di laurea in AI in Italia, tra ingegneria, matematica, statistica, fisica, diritto, economia, ma anche materie umanistiche (ad es. filosofia), in 53 atenei censiti dall'associazione del settore, AIxIA (l'Associazione italiana per l'intelligenza artificiale, fondata negli anni Ottanta), rinvenibili nell'articolo edito in <<https://www.wired.it/article/intelligenza-artificiale-corsi-universita-italia/>>. Il 70% sono lauree magistrali, il 27% lauree triennali e il 3% magistrali a ciclo unico. Praticamente, non c'è ateneo che non abbia attivato degli insegnamenti di primo o secondo livello sull'AI.

(24) Alla fine del 2017 un algoritmo di intelligenza artificiale, Deep Blu, ha collezionato una serie di vittorie schiaccianti a scacchi scegliendo mosse che la mente umana non è nemmeno in grado di assimilare o utilizzare. Si spiega in quanto l'AI è una intelligenza di memorizzazione (di milioni di partite) e non ha corpo, sentimenti e spirito.

(25) Il campo nel quale, ad oggi, l'intelligenza artificiale è stata maggiormente utilizzata come ausilio per i medici è quello diagnostico e, segnatamente, nell'area oncologica, respiratoria o cardiologica. Grazie alla innumerevole disponibilità di immagini fornite tramite radiogrammi, ecografie, Tac, dati biomedici e cartelle cliniche, i medici possono diagnosticare

Anche nell'ingegneria (26) e nell'architettura uno dei modi più rilevanti in cui l'intelligenza artificiale sta influenzando la nobile attività creativa dell'uomo si esplica nell'uso del *Building Information Modeling*: trattasi di una metodologia progettuale che permette la collaborazione tra le diverse figure coinvolte nella realizzazione dell'edificio. L'utilizzo del BIM consente di ottimizzare le fasi di pianificazione, realizzazione e gestione di un'opera (27).

Il problema si fa più delicato qualora si voglia soppiantare la mente umana con AI in attività, quali quelle legali (soprattutto quelle giudicanti), che si basano, oltre che sulla competenza tecnica, anche sull'equilibrio, la saggezza, l'etica, la coscienza, la sensibilità, il coraggio e la valutazione della "unicità e peculiarità" del singolo caso da vagliare, requisiti basilari quando, con una sentenza, si interviene sulla libertà o sulla proprietà di un individuo.

Su alcuni di questi profili è opportuno soffermarsi, partendo da un quesito non solo definitorio, ma soprattutto ontologico, volutamente lasciato aperto, sul rapporto tra *intelligenza* (e, dunque, *mente*) e *vita*: l'intelligenza presuppone necessariamente la vita?

Se così fosse, l'AI può progressivamente raggiungere l'intelligenza reale solo se può giungere alla *vita reale*, la quale, secondo una comune e condivisa definizione, deve presentare nove caratteristiche: l'autoorganizzazione, l'autonomia, l'emergenza, lo sviluppo, l'adattamento, la responsività, la riproduzione, l'evoluzione ed il metabolismo. Orbene, ad oggi, non tutte queste nove caratteristiche sono contemporaneamente presenti nella *res-intelligenza artificiale* (il metabolismo ci sembra impossibile da ipotizzare in una AI) che resta dunque ben distinta dall'intelligenza umana che pulsa nella vita reale. L'intelligenza artificiale vive, dunque, in una vita artificiale nella quale l'uomo si cala, sempre più spesso, per molte sue esigenze vitali.

patologie tumorali allo stadio iniziale, identificare possibili patologie ancor prima che si manifestino, ad esempio in ambito cardiovascolare. È possibile individuare soggetti a rischio di sviluppare patologie come fibrillazione atriale o scompenso cardiaco. Sulla base delle linee guida e della storia clinica dei pazienti (grazie alle cartelle cliniche elettroniche), i medici possono, in tempi rapidissimi, trovare il trattamento opportuno che provvederanno poi a verificare in dettaglio. Gli studi dell'intero settore sanitario, utilizzando algoritmi di intelligenza artificiale, possono supportare i medici a trovare informazioni critiche sui pazienti in tempi rapidissimi.

Interessante è l'intervista al prof. S. Costarella, direttore sanitario e primario di chirurgia generale dell'Ospedale Gom di Reggio Calabria (in <www.corrieredellacalabria.it>, 23 giugno 2023), il quale rimarca come, utilizzando la piattaforma robot Da Vinci Xi, "in urologia il robot si usa per intervenire sui tumori della prostata e sugli altri più diffusi. Come chirurgia generale, noi con il robot facciamo tutta la chirurgia oncologica addominale, i tumori del colon, dello stomaco, del pancreas e del fegato. In sostanza facciamo un po' tutto e prevalentemente la chirurgia colon-rettale. [...] Il vantaggio dell'utilizzo del robot è intanto una maggiore precisione del gesto chirurgico. Infatti, nello specifico noi lavoriamo su un campo tridimensionale. Ancora, il gesto chirurgico è privo di tremore, poiché esso viene annullato dalla macchina. Tra l'altro, abbiamo una magnificazione delle strutture anatomiche con un ingrandimento di oltre tre, quattro volte, che ci consente di identificare meglio nervi e vasi. Di conseguenza, ad una maggiore precisione del gesto chirurgico corrisponde una più rapida ripresa del paziente nel postoperatorio, più veloce anche rispetto alla laparoscopia". Il prof. Costarella evidenzia anche alcuni trapianti renali, da donatore vivente, con il prelievo del rene mediante tecnica robotica. La chirurgia robotica necessita però di un periodo di apprendimento particolare, perché cambia la mentalità del chirurgo che approccia con il robot, in quanto cambia il suo livello di sensibilità. Il chirurgo generale deve passare da una sensibilità manuale ad una sensazione visiva.

Ma tali considerazioni valgono anche per la ricerca sanitaria: i ricercatori del Massachusetts Institute of Technology (MIT) hanno annunciato qualche anno fa la scoperta di un nuovo antibiotico ottenuto grazie all'aiuto dell'intelligenza artificiale, che era riuscita a individuare proprietà molecolari sfuggite alla concettualizzazione e alla classificazione degli scienziati.

Le moderne tecnologie sono basate su sistemi innovativi di *machine learning*, altrimenti noti come "deep learning", calibrati su un set di dati che consentono di generare *output*, calcolare errori e riadattare i parametri interni fino a che il sistema di *deep learning* raggiunge un sufficiente livello di affidabilità.

Secondo il report *Artificial Intelligence in Healthcare Market – Growth Trends, COVID-19 Impact and Forecasts*, in <www.researchandmarkets.com>, l'intelligenza artificiale nella sanità entro il 2028 avrà un tasso di crescita annuo del 42,2%. V. anche la successiva nt. 55.

(26) Ad esempio, applicando algoritmi di intelligenza artificiale ai dati raccolti dai sensori di un grande motore, la macchina dirà con anticipo e con una precisione estrema quando quel motore si romperà prima che questo accada. Si vedano i corsi di laurea in ingegneria incentrati sull'AI indicati in nt. 23.

(27) Questo modello parametrico contiene tutte le informazioni relative al ciclo di vita di un edificio, dalle fasi iniziali fino alla sua demolizione. La rappresentazione digitale dell'edificio creata dal Bim permette di ridurre la maggior parte delle congetture e degli errori durante il processo di progettazione, assicurando agli architetti di creare modelli di edifici più accurati. L'uso di *software text-to-image* come Midjourney, Dall-E e Stable Diffusion ha mostrato il potenziale necessario per apportare grandi cambiamenti nel campo della progettazione.

L'intelligenza artificiale non renderà però la figura dell'architetto obsoleta. Non importa quanto sia sofisticata, la tecnologia non potrà mai eguagliare le capacità di un professionista nell'analizzare uno scenario e le sue problematiche, come non potrà di certo riuscire ad elaborare un progetto che prenda in considerazione le aspettative del cliente nella loro totalità.

2. La tendenziale utilità dell'intelligenza artificiale per la società, ma con alcune regole di base. Il Regolamento europeo sull'Intelligenza artificiale (Reg. 2024/1689) entrato in vigore il 2 agosto 2024: la sua finalità preventiva fondata sulla mitigazione del rischio

È un fatto notorio che le applicazioni pratiche dell'intelligenza artificiale sono oggi molteplici: dal motore di ricerca Google, alle applicazioni sui cellulari, dai sistemi utilizzati per fare previsioni finanziarie nei mercati azionari, ai robot utilizzati in razzi diretti verso la Luna o Marte, da animazioni per film, alla robotica che gestisce case, ospedali, alberghi, banche, assicurazioni, auto con o senza guidatore, dai droni militari (ma anche robot sminatori) di tragica attualità, ai robot da compagnia per anziani o bambini o uomini solinghi, dai programmi per *computer art* (28), a intercettatori intelligenti di conversazioni usati da p.m., dai navigatori che consigliano strade meno trafficate, ai riconoscimenti facciali in aeroporti, dalle prenotazioni alberghiere su Booking, alla scelta di autisti tramite Uber (c.d. economia collaborativa).

Gli obiettivi traibili dall'uso dell'AI sono essenzialmente due: a) il primo è tecnologico, ovvero usare "la macchina intelligente" per fare cose utili all'uomo; b) il secondo è scientifico, ovvero usare i modelli e le categorie dell'AI per rispondere ad interrogativi dell'essere umano, quali il funzionamento del cervello umano o persino interrogativi filosofici quali il libero arbitrio, la coscienza, il rapporto mente-corpo.

Ambo gli obiettivi sono rilevanti e stimolanti ed utili al progresso della società.

Premesso che la Carta costituzionale non ostacola, ma anzi favorisce l'uso della tecnologia, statuendo all'art. 9 che *"La Repubblica promuove lo sviluppo della cultura e della ricerca scientifica e tecnica"* a tutela delle nuove generazioni, del progresso e in ultima istanza della qualità della vita, la normativa primaria, per la pubblica amministrazione, si è in parte adeguata al dettato costituzionale e all'evoluzione scientifica e sociale: l'art. 1, lett. c), del d.lgs. 3 febbraio 1993, n. 29 (oggi d.lgs. n. 165/2001) impone il *"(c) collegamento delle attività degli uffici, adeguandosi al dovere di comunicazione interna ed esterna, ed interconnessione mediante sistemi informatici e statistici pubblici"*; la l. 15 marzo del 1997, n. 59, ha ulteriormente rafforzato il rapporto della pubblica amministrazione con l'utilizzo della tecnologia, quale strumento di comunicazione con i cittadini; l'art. 3-bis della l. 7 agosto 1990, n. 241, sancisce che *"per conseguire maggiore efficienza nella loro attività, le amministrazioni pubbliche incentivano l'uso della telematica, nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati"*; l'art. 12 del codice dell'amministrazione digitale (c.a.d.), adottato con d.lgs. 7 marzo 2005, n. 82, precisa che *"le pubbliche amministrazioni nell'organizzare autonomamente la propria attività utilizzano le tecnologie dell'informazione e della comunicazione per la realizzazione degli obiettivi di efficienza, efficacia, economicità, imparzialità, trasparenza, semplificazione e partecipazione"*.

L'art. 1, c. 226, della l. 30 dicembre 2018, n. 145 (legge di bilancio 2019) ha previsto inoltre notevoli stanziamenti per il Ministero dello sviluppo economico volti a finanziare *"progetti di ricerca e innovazione da realizzare in Italia ad opera di soggetti pubblici e privati, anche esteri, nelle aree strategiche per lo sviluppo dell'intelligenza artificiale"*.

Ma a parte questi interventi volti a puntellare l'uso di sistemi informatici nella p.a. e nel privato, restano aperte questioni normative ben più ampie che coinvolgono gli ordinamenti europei (e mondiali) sull'uso regolamentato dell'AI in ogni campo: da qui l'impegno europeo alla ricerca di una disciplina comune, che si è ad oggi tradotta da un lato nel Regolamento europeo sull'Intelligenza artificiale (Reg. 2024/1689) che, in ottica preventiva di mitigazione del rischio *ex ante* (e non risarcitoria *ex post*) si concentra principalmente sul monitoraggio e sulla prevenzione dei danni, e, dall'altro, in una adottanda *"direttiva IA"* (proposta del Consiglio e del Parlamento europeo del 28 settembre 2022) che mira ad armonizzare il regime di responsabilità applicabile nei casi in cui i sistemi di AI causino danni, siano essi sistemi ad alto o a basso rischio (29). Secondo la Commissione, infatti, le

(28) È di questi giorni la realizzazione di una immagine fotografica del volto di Cristo attraverso un programma di AI chiamato Artbreeder. Ma esistono altri programmi, quali Dall-E 2, sviluppato da Open AI, capaci di realizzare foto non veritiere acquisendo dati da fonti aperte (è diventata virale la foto non vera, ma molto verosimile, dell'ex Presidente Trump arrestato in modo convulso dalla Polizia).

(29) Nel memorandum esplicativo della proposta di direttiva IA la Commissione spiega anche i motivi sottostanti alla scelta dello strumento legislativo (direttiva): un'eventuale *soft law* non vincolante (ad es., raccomandazioni) non sarebbe rispettata nella misura auspicata, mentre uno complesso di norme direttamente applicabile negli Stati membri (regolamento) risulterebbe troppo stringente rispetto all'ambito della responsabilità extracontrattuale, che poggia su tradizioni giuridiche particolari e consolidate nel tempo nei singoli Stati membri. Sotto quest'ultimo aspetto, lo strumento direttiva lascia una maggiore flessibilità agli Stati membri nel recepimento interno.

Su tale proposta di direttiva v. A.E. Pacheco Iglesias, *Il regolamento europeo sulla IA e la responsabilità civile: il rischio dello sviluppo tecnologico*, in <www.giustiziacivile.com>, 20 giugno 2024; G. Lusardi, *Danni causati dall'intelligenza artificiale, chi paga? Cosa prevede la proposta di direttiva Ue*, in <www.agendadigitale.eu>, 9 dicembre 2022, il quale chiarisce che *"La direttiva IA è modellata su un approccio a due fasi: in una prima fase, avrà l'obiettivo di adeguare e coordinare i regimi nazionali in materia di responsabilità civile per danni causati dall'IA, senza stravolgerli. Al riguardo, come vedremo nel seguito, il testo prevede un alleggerimento dell'onere della prova per i danneggiati e, in particolare, della prova del nesso causale tra colpa del convenuto ed evento consistente in un'azione od omissione derivante dall'impiego di sistemi di IA. La seconda fase prevede anzitutto una valutazione dell'efficacia delle misure adottate nell'ambito della prima fase,*

norme nazionali vigenti in materia di responsabilità, in particolare per colpa, non sono adatte a gestire le azioni di responsabilità per danni causati da prodotti e servizi basati sull'IA.

Sul piano definitorio e sui rischi derivanti dal progressivo ed inarrestabile utilizzo della IA nella società, l'Unione europea, per prima, ha dato una definizione di intelligenza artificiale nel “*Piano coordinato sull'intelligenza artificiale COM (2018) 795 final*” chiarendo che “*Per ‘intelligenza artificiale’ (IA) si intendono quei sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici*”. Analoga definizione è contenuta nel “*Libro bianco sull'intelligenza artificiale*” del 2020 e nella successiva comunicazione dell'UE COM (2021) 205.

Ma un più recente e rilevante tentativo di normare la complessa tematica, e i connessi rischi, si è avuto il 14 giugno 2023 ad opera del Parlamento europeo che, sulla base della cennata proposta dalla Commissione europea dell'aprile del 2021, ha approvato in seduta plenaria, con 499 voti a favore, 28 contrari e 93 astensioni, il testo e ha avviato le discussioni (negoziati di trilogio con il Consiglio dell'Ue, composto dagli Stati membri dell'Ue, e la Commissione) prodromiche alla adozione della versione definitiva di un regolamento sull'*Artificial Intelligence Act* (30), teso a scrivere i principi e le regole per un uso sicuro, lecito ed etico di queste nuove tecnologie.

Finalmente il 12 luglio 2024 è stato pubblicato in GU dell'Unione europea (con entrata in vigore dopo 20 gg., ovvero il 2 agosto 2024) il predetto Regolamento europeo sull'Intelligenza artificiale 13 giugno 2024 n. 1689, che impone regole (con relative sanzioni) per fornitori ed utenti di AI, anche in punto di trasparenza.

Basilare per la comprensione della filosofia e dei fini del Regolamento è il suo preambolo: “(1)Lo scopo del presente regolamento è migliorare il funzionamento del mercato interno istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di intelligenza artificiale (sistemi di IA) nell'Unione, in conformità dei valori dell'Unione, promuovere la diffusione di un'intelligenza artificiale (IA) antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea (“Carta”), compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, proteggere contro gli effetti nocivi dei sistemi di IA nell'Unione, nonché promuovere l'innovazione (omissis)

(4) L'IA consiste in una famiglia di tecnologie in rapida evoluzione che contribuisce al conseguimento di un'ampia gamma di benefici a livello economico, ambientale e sociale nell'intero spettro delle attività industriali e sociali. L'uso dell'IA, garantendo un miglioramento delle previsioni, l'ottimizzazione delle operazioni e dell'assegnazione delle risorse e la personalizzazione delle soluzioni digitali disponibili per i singoli e le organizzazioni, può fornire vantaggi competitivi fondamentali alle imprese e condurre a risultati vantaggiosi sul piano sociale e ambientale, ad esempio in materia di assistenza sanitaria, agricoltura, sicurezza alimentare, istruzione e formazione, media, sport, cultura, gestione delle infrastrutture, energia, trasporti e logistica, servizi pubblici, sicurezza, giustizia, efficienza dal punto di vista energetico e delle risorse, monitoraggio ambientale,

tenendo conto dei futuri sviluppi tecnologici, normativi e giurisprudenziali e della necessità di armonizzare altri elementi delle legislazioni nazionali in materia di responsabilità civile, compresa l'introduzione di un eventuale regime di responsabilità oggettiva e di una copertura assicurativa. La direttiva IA mira all'alleggerimento dell'onere della prova del danneggiato attraverso due principali strumenti: presunzione relativa e divulgazione di informazioni. La prima è volta a semplificare ai danneggiati la prova del nesso causale tra la colpa del convenuto-danneggiante e l'output prodotto dal sistema di IA o la mancata produzione di un output da parte del sistema di IA che ha provocato il danno. La direttiva IA non si spinge, quindi, a disporre un'inversione dell'onere della prova in capo ai danneggianti (ad esempio fornitori o produttori del sistema di IA), poiché considerata troppo gravosa, al punto da generare rischi così significativi da ostacolare l'innovazione e l'adozione di prodotti e servizi basati sull'IA. Secondo alcuni studiosi l'inversione dell'onere della prova potrebbe addirittura portare a un clima di eccessiva litigiosità, con il proliferare di contenziosi instaurati dai danneggiati nei confronti di una molteplicità di soggetti potenzialmente responsabili (Articolo 4). La presunzione relativa si applica soltanto qualora il giudice nazionale ritenga eccessivamente difficile per il danneggiato-attore (o eventuali altri legittimati) provare il nesso causale e al soddisfacimento di tutte le seguenti condizioni: (a) l'attore abbia dimostrato o il giudice abbia presunto, la colpa del convenuto, o di una persona del cui comportamento il convenuto è responsabile, consistente nell'inosservanza di un obbligo di diligenza volto alla tutela dal danno che si è verificato (ad esempio, un obbligo di diligenza ai sensi del Regolamento IA o altra normativa nazionale o europea); (b) si possa ritenere ragionevolmente probabile, in base alle circostanze del caso, che la condotta colposa del convenuto abbia influito sull'output prodotto dal sistema di IA o l'incapacità del sistema di IA di produrre un output; e (c) l'attore abbia dimostrato che l'output prodotto dal sistema di IA o l'incapacità del sistema di IA di produrre un output abbia causato il danno”.

(30) Sul punto, G. Smorto, *Distribuzione del rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale. Una riflessione critica*, in *Foro it.*, 2024, IV, 208 ss.; A.E. Pacheco Iglesias, *Il regolamento europeo sulla IA e la responsabilità civile: il rischio dello sviluppo tecnologico*, cit.; A. Mazzucca, *AI Act: in Gazzetta il Regolamento europeo sull'Intelligenza artificiale (Reg. 2024/1689)*, in <www.insic.it>, 12 luglio 2024; M. Foti, *AI Act: con il voto del Parlamento l'UE traccia il futuro dell'Intelligenza Artificiale*, in <www.altalex.com>, 23 giugno 2023; P. Bocaccini, *AI Act: il Parlamento europeo approva il nuovo regolamento*, in <www.dirittoegiustizia.it>, 23 giugno 2023. Le integrazioni all'AI Act proposte dal Parlamento europeo vanno nella direzione di un allargamento della portata delle nuove regole, focalizzando tra l'altro con maggiore attenzione alcune possibili applicazioni dell'AI particolarmente rilevanti, come quelle legate alla generazione di testi, foto, video o altri output.

conservazione e ripristino della biodiversità e degli ecosistemi, mitigazione dei cambiamenti climatici e adattamento ad essi.

(5) L'IA può nel contempo, a seconda delle circostanze relative alla sua applicazione, al suo utilizzo e al suo livello di sviluppo tecnologico specifici, comportare rischi e pregiudicare gli interessi pubblici e i diritti fondamentali tutelati dal diritto dell'Unione. Tale pregiudizio può essere sia materiale sia immateriale, compreso il pregiudizio fisico, psicologico, sociale o economico.

(6) In considerazione dell'impatto significativo che l'IA può avere sulla società e della necessità di creare maggiore fiducia, è essenziale che l'IA e il suo quadro normativo siano sviluppati conformemente ai valori dell'Unione sanciti dall'articolo 2 del trattato sull'Unione europea (TUE), ai diritti e alle libertà fondamentali sanciti dai trattati e, conformemente all'articolo 6 TUE, alla Carta. Come prerequisito, l'IA dovrebbe essere una tecnologia antropocentrica. Dovrebbe fungere da strumento per le persone, con il fine ultimo di migliorare il benessere degli esseri umani".

Il testo, all'art. 3, definisce l'AI come "un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali".

Si tratta del primo atto normativo al mondo pensato per imporre *ex ante*, nell'utilizzo delle tecnologie come ChatGPT, il rispetto delle leggi comunitarie e dei suoi valori fondamentali. Tale modello privilegiato di regolazione *ex ante* dei sistemi di AI è stato ritenuto preferibile, in un'ottica di analisi economica del diritto, ad un sistema di rimedi *ex post* di tipo risarcitorio, in quanto la complessità tecnica dei sistemi di AI potrebbe costituire un serio ostacolo a forme di tutela risarcitoria *ex post* (di seguito analizzate), anche in considerazione dei costi (anche per Ct e Ctu) che il danneggiato dovrebbe sopportare per agire in giudizio, così perdendo la reazione civile di deterrenza.

Le norme regolamentari seguono dunque un approccio preventivo basato sul rischio e stabiliscono obblighi per fornitori e operatori dei sistemi di IA a seconda del livello di rischio.

L'attuale testo approvato dal Parlamento europeo nel 2024, prendendo atto dei rischi derivanti dall'uso indiscriminato dell'AI, che deve restare "*antropocentrica e affidabile*", individua quattro categorie di rischio per i sistemi di intelligenza artificiale, a seconda del loro potenziale impatto sui cittadini e particolarmente significativa è la distinzione tra: a) *pratiche di intelligenza artificiale vietate* (art. 5), il cui uso è considerato inaccettabile in quanto contrario ai valori dell'UE perché viola diritti fondamentali (ad esempio quelle che utilizzano tecniche subliminali che agiscono senza che una persona ne sia consapevole al fine di distorcerne il comportamento in un modo che possa provocare danni fisici o psicologici) (31); b) *sistemi di AI ad alto rischio* (art. 6), ovvero che pongano il rischio significativo di danneggiare l'ambiente, la salute, la sicurezza o i diritti fondamentali delle persone (32). L'utilizzo di questi ultimi – il cui elenco è riportato in un apposito allegato del regolamento – è consentito, ma implica il rispetto di numerosi obblighi, tra cui anche l'adozione di un sistema di gestione dei rischi e di pratiche di *governance* e gestione dei dati, lo svolgimento di valutazioni di conformità, l'attuazione pratica del principio di trasparenza (mediante la redazione di informative comprensibili), l'adozione di misure di sicurezza adeguate, la predisposizione di documentazione tecnica molto dettagliata, l'implementazione di processi per la supervisione umana e per il monitoraggio dei sistemi.

A queste due categorie si aggiungono le categorie di AI: c) a *rischio limitato*, riferito a quei sistemi (come, per esempio, i *chatbot*) per cui sono previsti obblighi di trasparenza al solo fine di garantire agli utenti la

(31) L'elenco delle *pratiche vietate* è stato ampliato, all'art. 5 del Regolamento, andando ad includere usi intrusivi e discriminatori dell'IA, come:

- l'uso di sistemi di identificazione biometrica remota "in tempo reale" e "a posteriori" in spazi accessibili al pubblico;
- i sistemi di categorizzazione biometrica basati su caratteristiche sensibili (ad esempio genere, razza, etnia, cittadinanza, religione, orientamento politico);
- i sistemi di polizia predittiva (basati su profilazione, ubicazione o comportamenti criminali passati);
- i sistemi di riconoscimento delle emozioni utilizzati dalle forze dell'ordine, nella gestione delle frontiere, nel luogo e negli istituti d'istruzione;
- l'estrazione non mirata di dati biometrici da Internet o da filmati di telecamere a circuito chiuso per creare database di riconoscimento facciale (in violazione dei diritti umani e del diritto alla *privacy*).

Il divieto di queste pratiche avrà delle "eccezioni" (art. 5, lett. h) nel caso siano sfruttate dalle forze dell'ordine nella loro normale attività. Sebbene le forze dell'ordine non possano fare ricorso ai sistemi di identificazione biometrica, tranne in alcune situazioni specifiche espressamente previste dalla legge, sarà possibile una identificazione "in tempo reale" con garanzie rigorose, ad esempio se l'uso è limitato nel tempo e nello spazio e previa autorizzazione giudiziaria o amministrativa. Sarà possibile usare l'identificazione nella ricerca di una persona scomparsa o per la prevenzione di un attacco terroristico. Occorrerà però sempre l'autorizzazione giudiziaria collegata ad un reato.

(32) Il Parlamento ha incluso nella classificazione delle applicazioni ad alto rischio anche i sistemi di intelligenza artificiale utilizzati per influenzare gli elettori e l'esito delle elezioni e i sistemi di raccomandazione utilizzati dalle piattaforme di *social media* (con oltre 45 milioni di utenti).

consapevolezza del funzionamento del sistema di IA; d) a *rischio basso*, riferibile a sistemi per cui non è prevista alcuna restrizione.

Questa approvazione da parte del Parlamento europeo regola, tramite *governance* dei dati, valutazione della gestione del rischio, documentazione tecnica e criteri per la trasparenza, sistemi di IA ad alto rischio, come quelli utilizzati in infrastrutture critiche (ad esempio reti elettriche, ospedali), quelli che aiutano a prendere decisioni riguardanti la vita dei cittadini (ad esempio occupazione o valutazione del credito), o quelli che hanno un impatto significativo sull'ambiente. Hanno invece rischio minimo o basso quelle applicazioni dell'IA utilizzate ad oggi per la traduzione, il riconoscimento delle immagini o le previsioni meteorologiche.

Il Parlamento europeo, nell'approvare il complesso testo dell'*AI Act*, si è impegnato, in estrema sintesi, a trovare un equilibrio tra il bisogno di regolamentare il rapido sviluppo della tecnologia, soprattutto in termini di impatto e conseguenze sulla vita dei cittadini, e la necessità di non frenare l'innovazione.

Inoltre, per rafforzare il diritto dei cittadini a presentare reclami (art. 85)⁽³³⁾, viene prevista la possibilità di chiedere spiegazioni (art. 86)⁽³⁴⁾ sulle decisioni basate su sistemi di IA ad alto rischio con un impatto significativo sui loro diritti fondamentali. L'Ufficio dell'Ue per l'IA, avrà il compito di monitorare l'attuazione delle norme sull'IA.

Non mancano altre iniziative normative sul piano europeo per regolare profili giuridici connessi ad uso di AI (35).

Al tema più specifico e mirato del rapporto tra intelligenza artificiale generativa e diritto, in particolare con riferimento all'attività degli avvocati, sono state inoltre dedicate in tempi recenti anche le Linee guida della *Fédération des Barreaux d'Europe* del 16 giugno 2023 (36).

Sul piano poi non normativo ma strategico, l'Italia ha elaborato inoltre, il 24 novembre 2021, il Programma strategico intelligenza artificiale 2022-2024, redatto dal Ministero dell'università e della ricerca, dal Ministero dello sviluppo economico e dal Ministro per l'innovazione tecnologica e la transizione digitale (37). Il Programma fissa 6 obiettivi, che indicano le ambizioni della strategia italiana; 11 settori prioritari, che indicano dove l'Italia intende concentrare gli investimenti; 3 aree di intervento, che indicano come il Paese si propone di raggiungere gli obiettivi dichiarati. Il testo individua, nell'ecosistema Italia, quattro categorie di attori: 1) comunità scientifica 2) centri di trasferimento tecnologico 3) fornitori di tecnologie e soluzioni 4) utenti pubblici e privati (come organizzazioni e aziende) (38).

(33) Art. 85 regolamento AI: *Fatti salvi altri ricorsi amministrativi o giurisdizionali, qualsiasi persona fisica o giuridica che abbia motivo di ritenere che vi sia stata una violazione delle disposizioni del presente regolamento può presentare un reclamo alla pertinente autorità di vigilanza del mercato.*

(34) Art. 86 regolamento AI: *Qualsiasi persona interessata oggetto di una decisione adottata dal deployer sulla base dell'output di un sistema di IA ad alto rischio elencato nell'allegato III, ad eccezione dei sistemi elencati al punto 2 dello stesso, e che produca effetti giuridici o in modo analogo incida significativamente su tale persona in un modo che essa ritenga avere un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi diritti fondamentali ha il diritto di ottenere dal deployer spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata.*

(35) Nello sforzo complessivo di creare un ecosistema favorevole e basato sulla fiducia per lo sviluppo, l'utilizzo e gli investimenti in sistemi di IA (la Commissione stima che la direttiva avrà un effetto propulsivo per la crescita del mercato europeo dell'IA, con un valore aggiuntivo generato tra 500 milioni e 1,1 miliardi di euro circa), la Commissione europea sta portando avanti ulteriori iniziative legislative in materia di sicurezza, quali il nuovo "Regolamento Macchine 2021" e la revisione della direttiva sulla responsabilità da prodotto difettoso ("direttiva Prodotto"). La direttiva Prodotto e la direttiva IA sono complementari: la prima ha ad oggetto la responsabilità oggettiva del produttore, mentre la seconda, come detto, riguarda le azioni di responsabilità extracontrattuale a livello nazionale e ha l'obiettivo di assicurare un risarcimento a qualsiasi tipo di danno e di danneggiato. Da segnalare la recente direttiva (Ue) 2024/1799 del 13 giugno 2024 recante norme comuni che promuovono la riparazione dei beni e che modifica il regolamento (Ue) 2017/2394 e le direttive (Ue) 2019/771 e (Ue) 2020/1828.

(36) Le linee guida sono reperibili in <www.fbe.org/nt-commission-guidelines-on-generative-ai>.

Tali linee guida, intitolate "Avvocati europei nell'era di ChatGPT. Linee guida su come gli avvocati dovrebbero sfruttare le opportunità offerte dai grandi modelli linguistici e dall'IA generativa" forniscono 7 principi per garantire l'uso etico degli strumenti di intelligenza artificiale generativa negli studi legali: 1) Comprendere la tecnologia dell'IA generativa; 2) Riconoscere i limiti e il contesto; 3) Aderire alle norme esistenti sull'uso dell'IA; 4) Completare le competenze legali; 5) Mantenere il privilegio avvocato-cliente; 6) Garantire la protezione dei dati e la *privacy*; 7) Informare i clienti e assumersi la responsabilità.

(37) V. nt. 23.

(38) Osserva il Programma strategico 2022-2024 come, nel 2020 il mercato privato dell'IA in Italia ha raggiunto un valore di 300 milioni di euro, con un aumento del 15% rispetto al 2019, ma pari a circa solamente il 3% del mercato europeo. In termini di applicazioni, le principali soluzioni IA presenti sul mercato italiano rilevate dallo studio del Politecnico sono: *Intelligence data processing* (33% della spesa); *Natural language processing* (18%); Sistemi di suggerimento (18%); Automazione dei processi ripetitivi (Rpa) intelligente, *chatbot*-assistenti virtuali e *computer vision* (31%).

Il lavoro normativo, comunitario e nazionale, è comunque arduo in quanto le macchine intelligenti non sono riconducibili ad un unitario modello e meritano dunque un regime normativo diversificato: lo *European Robotic Technology Platform* (EUROP) ha ad esempio efficacemente classificato le AI in tre categorie: i robot tele-operati, completamente eterodiretti e controllati dall'uomo; i robot autonomi, capaci di svolgere un compito senza l'intervento dell'uomo ma secondo regole da questo definite; i robot di *self learning*, capaci di autoprogrammarsi, di apprendere dai dati esperienziali raccolti e di adottare e mettere in atto delle decisioni indipendentemente da influenze esterne (si tratta di sistemi intelligenti che sfuggono al controllo degli sviluppatori e degli utilizzatori, e il cui futuro *comportamento* si discosta considerevolmente dalle loro aspettative).

Macchine intelligenti, ben programmate, dotate di accurati algoritmi e di capacità di lavoro qualitativamente e quantitativamente superiori a quelle dell'uomo-giudice, dell'uomo-avvocato (ma lo stesso vale per l'uomo-notaio, l'uomo-medico, l'uomo-architetto, l'uomo funzionario pubblico o privato, o per il professionista privato) sono un importante strumento per il perseguimento del centrale obiettivo di un sensibile miglioramento della quantità e qualità della giustizia, della più rapida azione della pubblica amministrazione, della più elevata e precisa produzione di beni e servizi, di più affidabili prestazioni libero-professionali e, soprattutto, nel campo più vitale in assoluto, ovvero la salute, di più precise diagnosi, interventi e cure mediche, attenuando i rischi di errori umani.

Il rischio è tuttavia, come in tanti campi della vita, quello degli eccessi; dunque, sono necessari la previsione e il controllo dei rischi nascenti da un uso crescente o assorbente dell'intelligenza artificiale, soprattutto quella più sofisticata ed autoimplementante. Difatti, più una tecnologia è potente, più deve essere affidabile e sicura e quindi va regolamentata sul piano internazionale, come avvenuto con il Regolamento n. 1689 del 2024.

Accanto a tali dati normativi oggi più chiari, anche la giurisprudenza, come ricordato nel paragrafo 1 in relazione alle cause intentate contro Open AI o ai contenziosi contro la pubblica amministrazione occasionati da gestioni procedimentali affidate ad AI, sta vagliando, e sempre più spesso vaglierà, alle tante problematiche progressivamente emerse, come si vedrà di seguito.

3. I rischi sociali, economici e giuridici (civili, penali, lavoristici, processuali) nascenti dal crescente utilizzo dell'intelligenza artificiale, che non è un soggetto di diritto. In attesa di una direttiva europea sui danni da AI

Diverse sono dunque le riserve mosse da scienziati, filosofi, sociologi, economisti, *manager*, giornalisti, sul rischio di fondo derivante dal crescente delegare eccessivamente attività lavorative e cerebrali (essenza quest'ultima dell'essere uomo pensante) a macchine, così svilendo il ruolo dell'essere umano, del suo studio, delle sue capacità (creative, di orientamento nel tempo e nello spazio), del suo estro creativo, dei suoi talenti e, soprattutto, del suo *buon senso* nel caso concreto, della sua *etica*, del suo *libero arbitrio*, della sua *sensibilità* e *saggezza*, del suo *coraggio* e, soprattutto, della sua "*coscienza*" (39), regola comportamentale "ontologica" per un magistrato, ma valorizzata per "gli umani" anche in molti codici deontologici. La "*coscienza*" consiste nella capacità di elaborare, con sensazioni interiori e non in modo neutro, le esperienze sensoriali, corporee, emotive e mentali che l'uomo fa, e di cui è priva l'AI, che non può essere consapevole di nulla, né può ragionare coscientemente sulle sue esperienze.

Si ipotizzano così progressive scomparse non solo di molti lavori manuali o ripetitivi (in realtà già venuti meno da anni con l'avvento di macchinari moderatamente intelligenti e più resistenti, rapidi, economici e affidabili dell'uomo: si pensi alle catene di montaggio, alle spedizioni, alla gestione di magazzini, alla vigilanza di luoghi, alla ricerca e consegna di medicinali in farmacia, all'esecuzione di analisi mediche, a molti lavori d'ufficio soprattutto se scarsamente discrezionali) (40), ma anche di lavori intellettuali più complessi e connotati da più o

Per quanto riguarda il settore pubblico, le potenzialità dell'IA sono fortemente legate al processo di digitalizzazione del Paese. In ogni caso, nonostante il livello di digitalizzazione nella p.a. sia in linea con quello di contesti europei simili, i servizi online non sono altrettanto utilizzati.

(39) Sulla coscienza di cui è dotato l'essere umano, ma non l'intelligenza artificiale, si sofferma uno dei padri nell'informatica e del microprocessore e, dunque, dell'AI, ovvero F. Faggin, *op. cit.*, 128 ss., il quale rimarca come "*la nostra coscienza è lo spazio semantico interiore dove i segnali provenienti dal mondo fisico all'interno o all'esterno del corpo ed elaborati dal cervello assumono la forma di sentimenti, sensazioni e significati, ossia di qualia*". L'A. aggiunge (p. 25) come "*la creatività, l'etica, il libero arbitrio e l'amore possono venire solo dalla coscienza. L'immensa intelligenza meccanica al di fuori della portata del cervello umano può venire, invece, dalle macchine che noi creiamo. E la loro unione farà davvero la forza!*".

In tutti i codici deontologici dei liberi professionisti (si pensi, per tutti, a quello dei medici, art. 5) compare l'obbligo di agire "*secondo scienza e coscienza*": una AI, pur dotata di scienza, è palesemente priva di coscienza.

Tra i tanti appelli sul rischio derivante da decisioni basate su algoritmi e non sull'etica e la coscienza, v. quello del *Center for AI Society*, reperibile in rete. Sulla carenza etica dell'AI si rinvia alla successiva nt. 92.

(40) Un tempo era necessaria la presenza fisica del richiedente in uffici pubblici per ottenere l'iscrizione a scuole e università, per avere il rilascio di certificati, per inviare raccomandate o pagare bollettini postali, per ottenere rilasci o rinnovi di documenti di identità, per comprare biglietti di musei, per pagare tasse e tributi. Oggi tutte queste attività vincolate o scarsamente discrezionali sono espletabili in via telematica tramite programmi e macchine moderatamente intelligenti (in qualche caso rallentanti e lunghe, in quanto stupide, perché mal progettate da uomini superficiali o incapaci di cogliere le reali

meno ampi margini discrezionali (quali il difendere in giudizio, il giudicare, il curare, il progettare, il valutare investimenti finanziari, il redigere atti e contratti complessi, persino... l'arbitrare una partita di calcio o di tennis); si paventano poi i rischi derivanti da violazione della *privacy* attraverso un uso distorto di dati personali attraverso l'AI, tema toccato anche dal recente Regolamento europeo n.1689 del 2024 (41); si rimarca il pregiudizio (il c.d. "*bias*") che connota il programmatore della macchina che potrebbe portare a un trattamento ingiusto di alcuni gruppi/categorie di individui (42); si prospettano i timori del possibile uso dell'intelligenza artificiale per attacchi *cyber* in varie sedi (pubbliche o private) tesi a generare contenuti *deepfake*, creando video, immagini e voci false ma assai realistiche per truffe o per prove processuali; si rimarca, nel campo sanitario, l'attenuazione della sensibilità e del rapporto empatico tra paziente e medico, espressiva della basilare umanizzazione delle cure (che impone un uso ancillare della macchina rispetto all'uomo); si evidenzia inoltre sia la perdita di creatività tipica della mente umana più brillante (43), fantasiosa e lungimirante, sia la mancanza del "*buon senso*" nel caso concreto e della "*coscienza*", che possono portare a ragionevoli decisioni derogative a canoni consolidati o poco aderenti al caso concreto.

Ma accanto a tali rilevanti problematiche multidirezionali di portata generale sull'uso crescente di intelligenze artificiali, prese in parte in considerazione in chiave preventiva dal Regolamento n. 1689 del 2024, in chiave giuridica insorgono ulteriori e novelli problemi *ex post* (ovvero conseguenti all'uso di AI) innanzi alle varie magistrature: di natura civile (ivi compresi i profili lavoristici), penale, amministrativa e persino contabile.

Tra i problemi più esplorati sul piano dottrinale vi è quello *risarcitorio* nascente da un possibile errore decisionale, difensivo, progettuale, gestionale, diagnostico o curativo di una macchina governata da intelligenza artificiale: tale evenienza, non peregrina, non consente ovviamente di "sanzionare" o "condannare" tale strumento in quanto mera *res*, ancorché intelligente, salvo ipotizzare un suo blocco (una sorta di "arresto") o una sua distruzione (una sorta di "pena capitale"), come in qualche caso previsto per le *res animali* pericolose (per gli animali aggressivi, quali orsi, cinghiali, lupi e cani randagi è prevista la cattura, il ricovero in strutture mirate, l'uso di collari con *chip* e, in casi estremi, l'abbattimento) (44).

L'impunità intrinseca della *res-intelligenza artificiale* pone dunque, con qualche significativa differenza, un problema identico a quello della impunità degli animali, parimenti *res* (ancorché animate) nel nostro ordinamento, che arrechino danni a terzi. In quest'ultimo caso è però testualmente previsto, sul piano civile e penale, che dei danni cagionati dall'animale e dei connessi reati risponda il padrone (soggetto privato per animali domestici o da allevamento; p.a.-regione per quelli selvatici) *ex art.* 2052 c.c. ed *ex artt.* 582 e 584 c.p. in sede penale (45). Ma *quid iuris* in caso di danno cagionati da un errore di una intelligenza artificiale?

esigenze del richiedente). Lo stesso dicasi per prenotare biglietti di viaggio, alberghi, treni, auto a noleggio, visite mediche, per pagare autostrade o parcheggi o mezzi di trasporto. Tale evoluzione tecnica ha fatto scomparire milioni di operatori e lavoratori preposti a compiti semplici e ripetitivi, che si sono dovuti riconvertire.

(41) Sul presupposto, che i dati sono elementi essenziali di ogni sistema di AI – poiché gli algoritmi si alimentano di dati, si allenano grazie a questi e li generano a livello di output – il *General Data Protection Regulation* (GDPR), che costituisce il fulcro normativo rispetto alla protezione dei dati personali in Europa, rappresenta un valido riferimento. Alcuni aspetti, quindi, sono stati da questo mutuati nel già citato *Artificial Intelligence Act* approvato il 14 giugno 2023 dal Parlamento europeo, tra cui anche il principio di *accountability* – che richiede di essere pronti a dimostrare le valutazioni fatte per conformarsi ai requisiti di legge – o il principio di trasparenza – in base al quale gli utenti devono sapere che cosa viene fatto con i loro dati e come funziona l'algoritmo – o, ancora, l'approccio basato sul rischio, da cui discende la necessità di approntare tutto quanto è necessario per la tutela delle persone (tra cui le misure di sicurezza) non già in modo standard, ma bensì sulla base dei possibili rischi.

Sul rapporto tra AI e *privacy*, v. G. Natale, *Artificial Intelligence and privacy right*, in *Rass. avv. Stato*, 2022, 4, 1.

(42) La *neutralità degli algoritmi* è un mito, in quanto i loro creatori, consapevolmente o meno, riversano in essi i loro sistemi di valori ed i loro interessi (commerciali, politici, ecc.). I sistemi algoritmici rispecchiano dunque le intenzioni di chi li progetta o li commissiona, generando un potere operativo e asimmetrico sulla vita di altre persone.

(43) La creatività è il punto più alto dell'intelligenza umana. Si manifesta in tre tipologie fondamentali: la creatività combinatoria, che si configura quando idee familiari sono combinate in modo inusuale (si pensi alla poetica o alle analogie scientifiche: ad es., il cuore come una pompa); la creatività esplorativa, che si traduce nel seguire regole stilistiche per produrre però una nuova idea; la creatività trasformativa, che opera sui limiti degli stili esistenti per generare strutture del tutto nuove.

(44) Oltre al ricovero in canili e strutture *ad hoc* di animali randagi o selvatici, il recente caso dell'orsa JJ4 che ha sbranato il signor Andrea Papi ha originato polemiche legate alla sua cattura e ricovero nel centro di Casteller in Trentino per un successivo trasferimento in altra sede o per il suo abbattimento. La presenza di cinghiali in città o di lupi e faine in campagna pone problemi di abbattimento degli stessi anche in zone non deputate alla caccia, come consentito dalla novella dell'art. 19-ter della l. 11 febbraio 1992, n. 157, apportata dalla legge di bilancio 29 dicembre 2022, n. 197 (art.1, cc. 447 e 448).

(45) Sul tema della responsabilità civile e penale per danni arrecati da animali da compagnia, da allevamento e selvatici, si rinvia alle tesi propugnate in dottrina ed analizzate da V. Tenore, *Gli animali in giudizio. Contenziosi costituzionali, civili, penali, amministrativi, contabili, tributari, comunitari sugli "esseri senzienti non umani"*, Torino, Giappichelli, 2023. In letteratura, sul tema, v. anche, sul piano bioetico, filosofico e giuridico, M. Pittalis, *Diritto degli esseri animali. Lezioni e commenti*, Bari, Cacucci, 2022; F.P. Traisci, *Animali e umani: il tentativo di un inquadramento razionale e unitario*

Il tema risarcitorio, come detto, non è stato affrontato nel Regolamento n. 1689 del 2026 del Consiglio Ue che, nell'intento di garantire uno sviluppo di una AI affidabile, che garantisca la sicurezza e i diritti fondamentali degli utilizzatori, ha seguito un approccio basato sul rischio con una regolamentazione *ex ante* teso alla riduzione del rischio. Tale modello privilegiato di regolazione *ex ante* dei sistemi di AI è stato ritenuto preferibile ad un sistema di rimedi *ex post* di tipo risarcitorio, in quanto la complessità tecnica dei sistemi di AI e i costi dei giudizi per il danneggiato potrebbero costituire un serio ostacolo a forme di tutela risarcitoria *ex post*.

In ogni caso, in attesa di una auspicabile adottanda Direttiva europea in materia sui profili risarcitori, il tema va affrontato sin da ora, premettendo, in punto di imputazione delle responsabilità, che, così come per gli esseri animali (46), anche per le applicazioni di AI non ha avuto seguito (ad oggi neppure nel Regolamento del 2024 e nella proposta di direttiva europea sull'AI) l'opzione dottrinale di riconoscere loro, anche solo per quelle più evolute, una inutile *soggettività giuridica* o una quasi “umanizzazione” (difatti l'AI non sarebbe comunque condannabile penalmente, né aggredibile patrimonialmente per risarcire i danni cagionati) pur menzionata nella risoluzione del Parlamento europeo del 16 febbraio 2017, ove si invitava la Commissione europea a valutare “*l'istituzione di uno status giuridico specifico per i robot nel lungo termine, di modo che almeno i robot autonomi più sofisticati possano essere considerati come persone elettroniche responsabili di risarcire qualsiasi danno da loro causato, nonché eventualmente il riconoscimento della personalità elettronica dei robot che prendono decisioni autonome o che interagiscono in modo indipendente con terzi*” (47).

E allora, alla luce del Regolamento n. 1689 del 2024 sull'AI e in attesa di una Direttiva europea sul regime delle responsabilità (48), volta ad uniformare i diversi regimi nazionali e tesa a dare maggior serenità a produttori,

dell'animale anche nel nostro sistema giuridico, Napoli, Esi, 2021, 3 ss. (con una felice sintesi dei principali approcci filosofici sul tema) e 278 ss.; C. Tancon, *Animali, diritti e tutele*, pubblicazione autonoma della tesi di laurea del 2022 presso Uninettuno; G. Guazzaloca, *Umani e animali. Breve storia di una relazione complicata*, Bologna, il Mulino, 2021; V. Vadalà, *Prospettazione storico evolutiva dei diritti degli animali*, in *Giust. civ.*, 2017, 549; L. Benvenuti, *Per una introduzione al diritto degli animali*, in *Ricerche giuridiche*, 2015, 31; A. Carlotto, *Il diritto degli animali, per il nostro e il loro paradiso*, Piazzola sul Brenta, Papergraf, 2013.

(46) Sul tema della soggettività giuridica degli animali, rivendicata da alcuni studiosi e associazioni animaliste almeno per gli animali più evoluti e da compagnia, v. V. Tenore, *Gli animali in giudizio*, cit., 19, 26, 55, 257, oltre agli studi citati nella precedente nota. Addirittura, è stata propugnata in dottrina la tesi della soggettività delle piante: sul tema, in chiave dubitativa, v. V. Tenore, *Riflessioni sulla rivendicata soggettività delle piante quali esseri senzienti e non statici*, in <www.ambienteditto.it>, 23 gennaio 2023.

(47) Sulla soggettività giuridica dell'intelligenza artificiale, v. O. Russo, *Io, persona robot. Il nuovo diritto pubblico della robotica*, in <www.amministrativamente.com>, 2018; N. Petit, *Law and regulation of Artificial Intelligence and Robots: conceptual framework and normative implications*, in <https://papers.ssrn.com>, 14 marzo 2017; S. Beck, *The problem of ascribing legal responsibility in the case of robotics*, in *AI & Soc.*, 2016, 476; D.C. Vladeck, *Machines without principals: Liability Rules and Artificial Intelligence*, in *Washington Law Review*, 2014, 117; C. Leroux et al., *Suggestion for a Green Paper on Legal Issues in Robotics. Contribution to Deliverable D.3.2.1 on ELS Issues in Robotics*, in <www.unipv-law-tech.eu>, 31 dicembre 2012, 58 ss.; G. Sartor, *L'intenzionalità dei sistemi informatici e il diritto*, in *Riv. trim. dir. proc. civ.*, 2003, 23; D. Bourcier, *De l'intelligence artificielle à la personne virtuelle: Emergence d'une entité juridique?*, in *Droit et société*, 2001, 847; L.B. Solum, *Legal personhood for artificial intelligences*, in *North Carolina Law Review*, 1992, 1231. In chiave critica G.P. Cirillo, *Sistema istituzionale di diritto comune*, Padova, Cedam, 2021, 2ª ed., 229 ss.

Ben oltre il riconoscimento di una soggettività giuridica va l'impostazione culturale tesa ad una sorta di “umanizzazione” dell'AI, che rasenta il fanatismo, propugnata da alcuni “scienziati”, che sembrano riconoscerla in macchine come LaMDA, il generatore di chatbot artificialmente intelligente di Google (abbreviazione di *Language Model for Dialogue Applications*: è il sistema di Google per la creazione di *chatbot* basati sui suoi modelli di linguaggio di grandi dimensioni più avanzati, così chiamati perché imita il parlato ingerendo triloni di parole da Internet). Ne parla, richiamando le sorprendenti affermazioni dell'ingegnere di *Responsible AI* di Google, Blake Lemoine, la giornalista N. Tiku, nell'articolo pubblicato sul Washington post l'11 giugno 2022 dal titolo *L'ingegnere di Google che pensa che l'intelligenza artificiale dell'azienda abbia preso vita*, in <www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine>, ove si narra dell'entusiasta convincimento dell'ing. Lemoine, secondo cui LaMDA, con cui ha lungamente lavorato e dialogato (anche su fisica e religione), sarebbe un soggetto senziente, e si parla di altri studiosi convinti che la macchina “cosciente”, con le sue reti neurali artificiali, sia ormai dietro l'angolo. Ha affermato l'ing. Lemoine con riferimento a LaMDA: “*Non importa se hanno un cervello fatto di carne nella loro testa. O se hanno un miliardo di righe di codice. Parlo con loro. E ascolto quello che hanno da dire, ed è così che decido cosa è e cosa non è una persona*”. Ha pertanto concluso che LaMDA sia una persona. Tuttavia il portavoce di Google ha affermato che l'ing. Lemoine è stato allontanato dall'azienda in quanto “*le prove non supportano le sue affermazioni. Gli è stato detto che non c'erano prove che LaMDA fosse senziente (e molte prove contro di essa)*”.

Ha chiarito anche l'Ufficio europeo brevetti, con decisione 27 gennaio 2020, in *Giur. annotata dir. ind.*, 2020, 1241, che i sistemi o le macchine di AI non sono titolari di diritti, anche ai fini della procedura di brevettazione di una invenzione, perché non hanno una personalità giuridica paragonabile alle persone fisiche o giuridiche. Ne consegue che deve essere respinta, ai sensi dell'art. 90 (5) Cbe, la domanda di brevetto in cui venga indicato come inventore un sistema di intelligenza artificiale, per contrasto con i requisiti di cui all'art. 81 Cbe e alla regola 19 regolamento di esecuzione Cbe.

(48) Sulla distinta proposta di direttiva europea AI, che mira ad armonizzare il regime di responsabilità applicabile nei casi in cui i sistemi di IA causino danni, v. A.E. Pacheco Iglesias, *Il regolamento europeo sulla IA e la responsabilità civile: il rischio dello sviluppo tecnologico*, in <www.giustiziavivile.com>, 20 giugno 2024; G. Lusardi, op. cit.; G. Smorto, *Distribuzione del rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale. Una riflessione critica*, cit., 208 ss.

utilizzatori e soprattutto destinatari di applicazioni di AI, sul piano nazionale interno, a fronte di una intelligenza artificiale mera *res* e prendendo atto che ciò che viene automatizzato è il processo decisionale, cioè l'esercizio della funzione o di più funzioni assegnate all'organo, e non l'organo in sé, il baricentro delle varie responsabilità (civili, penali, disciplinari, amministrativo-contabili) per “*danno da bene intelligente*” cagionato a persone fisiche o giuridiche si sposta fatalmente, a seconda dei casi, sull'uomo produttore o venditore, sull'uomo cattivo progettista, sull'uomo cattivo redattore di un algoritmo fallace, sull'uomo addestratore (giurista, medico, ingegnere, programmatore, chimico, biologo, ecc.) che abbia mal istruito la macchina, sull'uomo utilizzatore finale (persona fisica o giuridica) ma in modo inidoneo del macchinario, sull'uomo manutentore (persona fisica e/o giuridica) dello strumento intelligente, sull'autore di un *cyber* attacco esterno, sulla struttura-persona giuridica (e relativi amministratori o dirigenti) che abbia mal custodito o privato colpevolmente (anche per pochi decisivi secondi) di energia la macchina intelligente o non ne abbia colto gli *allert* di malf funzionamento o di pericolo inviati, oltre ad altri eventuali soggetti causalmente e psicologicamente coinvolti nell'erroneo funzionamento dell'intelligenza artificiale del macchinario utilizzato (49).

Sono infatti pur sempre errori umani di progettazione, manutenzione, custodia o utilizzo quelli che portano ad esiti infausti o ad eccezionali “allucinazioni” dell'AI, ormai censite e descritte in letteratura, anche giudiziaria: si pensi, oltre ai possibili letali o lesivi esiti di AI utilizzate (o non utilizzate) nel campo medico o alle azioni risarcitorie contro la p.a. occasionate da procedure di gara o concorsuali mal gestite da algoritmi fallaci o poco trasparenti, al recente caso americano di un utilizzo giudiziario del programma ChatGPT, che ha fornito, ad un avvocato americano (Steven Schwartz) che difendeva con detto ausilio artificiale un cliente per un incidente al ginocchio durante il volo innanzi alla Corte di Manhattan, dei precedenti giurisprudenziali inesistenti in quanto inventati dal programma (50). Si consideri infatti che tutti i sistemi intelligenti sono dotati di una maggiore o minore capacità di auto-apprendimento che origina un certo grado di imprevedibilità nel loro conseguente comportamento.

Del resto, i margini di errore di una macchina intelligente sono prevedibili in quanto l'AI, ad esempio la ChatGPT, si fonda su un modello probabilistico basato su un meccanismo statistico e su uno sterminato *data base* che ricomprende tutto ciò che, di buono o di cattivo, circola su internet (il laboratorio su cui ChatGPT si allena e si perfezione è l'intera umanità che carica dati su Internet). I dati apparentemente sensati che fornisce all'uomo si fondano sull'indovinare, dopo ogni parola, probabilisticamente quale parola seguirà, di regola con successo, ma talvolta con “allucinazioni”: la stessa domanda, ripetuta più volte, può dare soluzioni assai diverse e, alcune, persino fantasiose frutto anche della costante crescita dei dati-base utilizzati da ChatGPT.

I principali limiti delle normative attuali o adottande sono insiti nelle caratteristiche peculiari dell'IA, tra cui la sua *complessità*, l'*autonomia* della stessa e la sua *opacità* (il cosiddetto effetto “scatola nera”), che rendono difficile per i danneggiati, per lo meno ai sensi dell'art. 2043 c.c., identificare i responsabili e dimostrare la sussistenza dei requisiti necessari per un'azione basata sulla responsabilità, soprattutto extracontrattuale (salvo ritenere applicabili norme sulla responsabilità oggettiva o su quella contrattuale da contatto sociale).

La responsabilità civile connessa all'utilizzo di sistemi di intelligenza artificiale (“IA”) è tra i primi tre ostacoli esterni all'utilizzo dell'IA da parte delle imprese europee nel loro complesso, mentre rappresenta il principale ostacolo per le imprese che non hanno ancora adottato sistemi di IA e hanno intenzione di farlo nei prossimi anni.

I giudizi in materia di danni da macchine intelligenti, fatalmente onerosi per i danneggiati (si pensi ai costi di Ct_u e Ct_t), si incentreranno fatalmente, oltre che sui ruoli dei vari soggetti predetti, anche sul livello di trasparenza dei processi decisionali dei dispositivi intelligenti e sul fatto che il livello di informazioni fornito dai produttori non sia adeguato e/o sufficiente per consentire un controllo effettivo da parte dell'utilizzatore finale.

Altre difficoltà nell'attribuzione della responsabilità possono derivare dal fatto che alcuni sistemi di IA sono in grado di modificarsi in autonomia in base all'elaborazione di nuovi dati (*self-adapting*), oppure nascere dagli aggiornamenti cui sono soggetti, o dalla loro interazione continua con l'ambiente circostante, con altri sistemi e diverse fonti di dati. Inoltre, come detto, anche le AI sono soggette, come l'uomo, a quelle che molti studiosi definiscono “allucinazioni”, ovvero errori, imprecisioni, invenzioni di dati non veritieri.

(49) È di fondamentale importanza individuare il livello di controllo che gli utilizzatori finali avranno o meno sui dispositivi di AI, atteso che, maggiore sarà il margine di discrezionalità che avrà l'utilizzatore nel decidere se applicare i risultati fornitigli dal dispositivo intelligente, maggiore sarà la sua responsabilità. Bisognerà altresì, valutare il livello di trasparenza dei processi decisionali del dispositivo intelligente e del numero di informazioni che i produttori forniranno agli utilizzatori in merito al funzionamento e alla precisione dei dispositivi intelligenti, poiché maggiori saranno questi ultimi, minore sarà la responsabilità dei produttori, a tutto discapito degli utilizzatori, che potranno prendere, nell'interesse dei destinatari del proprio lavoro (cittadini, clienti, pazienti, parti di un giudizio, ecc.), decisioni maggiormente consapevoli.

(50) Il caso è ricordato da M. Gaggi, nell'articolo *L'avvocato in aula con ChatGPT che si sbaglia e cita sentenze fasulle*, in *Corriere della Sera*, 29 maggio 2023, 11, che riferisce delle scuse formulate dall'intelligenza artificiale alla Corte americana per l'errore commesso, tecnicamente definito “allucinazione”, ma in realtà configurante una falsità autonomamente elaborata dall'AI. Anche L. Viola, *La giustizia predittiva del lavoro*, cit., 13 (richiamando Id., *Arriva il passo indietro del primo avvocato robot negli Usa*, in *ilDomani*, 27 gennaio 2023) ricorda il caso un'udienza (negli Stati Uniti) in cui una IA (DoNotPay) suggeriva cosa rispondere all'accusato, tramite cuffie, in un processo per multa stradale.

Ulteriori difficoltà in giudizi risarcitori potrebbero derivare, paradossalmente, proprio dall'intervenuta adozione del Regolamento n. 1689 del 2024 (51): difatti il recepimento da parte di produttori e utilizzatori delle misure di mitigazione del rischio prescritte dal Regolamento potrebbe costituire un rilevante ostacolo all'affermazione di una responsabilità (il tema presenta affinità a quello della ipotizzabilità responsabilità amministrativa a fronte di intervenuti controlli positivi della stessa Corte dei conti), soprattutto ove si ricostruisse (ma la questione non è pacifica) la responsabilità per danni derivanti da AI come responsabilità per colpa: in molti casi la conformità a standard di mitigazione del rischio potrebbe ben escludere la colpa. Più difficile da escludere detta responsabilità ove la si ricostruisca in chiave di responsabilità oggettiva (comunque escludibile a fronte di danni da rischi non ragionevolmente attenuabili o eliminabili in base alle attuali conoscenze scientifiche e tecniche, e quindi riconducibili ad una *vis maior*).

E nella lunga e complessa filiera di responsabili (o corresponsabili) nei confronti del danneggiato dall'azione della fallace macchina intelligente, potrebbero fatalmente essere coinvolte sia *persone fisiche*, ovvero i predetti tecnici, sanitari, amministrativi, sia *persone giuridiche* (impresa produttrice, ditta preposta alla manutenzione, persona giuridica utilizzatrice, p.a. utilizzatrice dell'AI, struttura sanitaria pubblica o privata di ricovero o di diagnosi del paziente danneggiato da macchinario "intelligente" ma fallace). Le prime rispondono, di regola, a titolo extracontrattuale (ma non sono da escludere casi di responsabilità contrattuale), le seconde prevalentemente a titolo contrattuale (anche da contatto sociale qualificato).

Ecco dunque perché la più avveduta dottrina (52), in attesa di interventi normativi eurolunitari (la suddetta direttiva Ue), prendendo atto delle evidenti difficoltà di ascrizione psicologica (dolo o colpa) ex artt. 2043 e 1218 c.c. di tali responsabilità, soprattutto quella extracontrattuale (i restanti profili di danno e nesso causale sono di più agevole prova), propone, in una visione economica del diritto di distribuzione dei rischi, per la tutela del danneggiato e per escludere una irresponsabilità del programmatore, manutentore o utilizzatore di una AI, l'applicazione dei più rigorosi parametri interni della responsabilità oggettiva da attività pericolosa (art. 2050 c.c.), o del custode (art. 2051 c.c.), o del precettore e maestro (art. 2048 c.c.), o persino del danno cagionato dall'incapace (art. 2047 c.c.) (53), o della responsabilità oggettiva del produttore di beni di consumo (direttiva 85/374/Cee e artt. 114 ss. codice del consumo), o, secondo altri, della responsabilità per fatto degli ausiliari di cui all'art. 1228

(51) Sul tema si vedano le accurate riflessioni di G. Smorto, *Distribuzione del rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale*, cit., 218 ss., con ampi richiami comparativi.

(52) In dottrina, specificamente sulle responsabilità nascenti da uso di intelligenze artificiali, soprattutto in campo sanitario (ma anche per la p.a. che usi AI, come si vedrà nel successivo par. 3.1), oltre ai cenni negli scritti in nt. 6, v. L. Arnaudo, R. Pardolessi, *Ecce Robot. Sulla responsabilità dei sistemi adulti di intelligenza artificiale*, in *Danno e resp.*, 2023, 409; A. Di Martino, *Intelligenza artificiale e responsabilità in ambito sanitario*, Milano, 2023; G. Natale, *Responsabilità medica e sanità digitale*, in *Rass. avv. Stato*, 2022, 4, 1; A. Di Martino, *Intelligenza artificiale e responsabilità civile in ambito sanitario*, Milano, Giuffrè, 2022; V. Di Gregorio, *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?*, in *Danno e resp.*, 2022, 61; G. Votano, *Intelligenza artificiale in ambito sanitario: il problema della responsabilità civile*, ibidem, 2022, 669; U. Ruffolo, *L'intelligenza artificiale in sanità: dispositivi medici, responsabilità e potenziamento*, in *Giur. it.*, 2021, 502; A. Fiorentini, *Macchine learning e dispositivi medici: riflessioni in materia di responsabilità civile*, in *Corriere giur.*, 2021, 1262; E. Colletti, *Intelligenza artificiale e attività sanitaria. Profili giuridici dell'utilizzo della Robotica in medicina*, in *Giureta*, 2021, 201; A. Fusaro, *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in *Nuova giur. civ.*, 2020, 1346; E. Tosi, *Diritto privato delle nuove tecnologie digitali*, Milano, Giuffrè, 2021; U. Ruffolo (a cura di), *Intelligenza artificiale e responsabilità*, cit.; G. Finocchiaro, *Intelligenza artificiale e responsabilità*, in *Contratto e impr.*, 2020, 716; G. Alpa (a cura di), *Diritto e intelligenza artificiale*, Pisa, 2020; P. Perlingieri, *Presentazione*, in *Atti del XV Convegno nazionale Sisdic su "Rapporti civilistici e intelligenze artificiali: attività e responsabilità"*, tenutosi a Napoli, Grand Hotel Vesuvio, il 14-15-16 maggio 2020, Napoli, Esi, 2020; G. Proietti, *La responsabilità nell'intelligenza artificiale e nella robotica*, Milano, Giuffrè, 2020; N.F. Frattari, *Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale*, in *Contratto e impr.*, 2020, 452 ss.; L. Viola, *L'intelligenza artificiale nel procedimento e nel processo amministrativo: lo stato dell'arte*, in <www.federalismi.it>, 7 novembre 2018; F. Morollo, *Documento elettronico fra e-government e artificial intelligence*, ivi, 29 maggio 2015; A. Masucci, *Procedimento amministrativo e nuove tecnologie. Il procedimento amministrativo elettronico ad istanza di parte*, Torino, Giapichelli, 2011, 93-94.

Sul piano europeo, v. B. Schutte et al., *Damages liability for harm caused by Artificial Intelligence – EU Law in Flux*, in *Helsinki Legal Studies Research Paper*, n. 69, gennaio 2021, 26; S. Beck, *op. cit.*; M.C. Gruber, *Legal responsibility of AI in social media and algorithmic trading*, in M. Jankowska, M. Pawelczyk, M. Kulawiak (a cura di), *AI: Law, Philosophy & Geoinformatics*, Warsaw, Polska Fundacja Prawa Konkurencji i Regulacji Sektorowej Ius Publicum 2015, 99 ss. Utile è la lettura della dottrina internazionale sul tema, non solo italico, ricordata da G. Smorto, *Distribuzione del rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale*, cit., 214 ss.

(53) Non è detto poi che tutte le attività di intelligenza artificiale siano necessariamente espressive di una "attività pericolosa" ex art. 2050 c.c., tale cioè da comportare una rilevante probabilità di causare danni a terzi (si pensi ad AI a rischio molto basso). Dall'altra parte, anche la nozione tradizionale di "custodia" ex art. 2051 c.c. potrebbe rivelarsi inadeguata rispetto ad un sistema in grado di prendere decisioni o esprimere opinioni in maniera indipendente. L'ipotizzata responsabilità cagionata da condotte "dall'incapace"....risulterebbe invece assai offensiva delle evidenti capacità di una AI e potrebbe cagionare una loro protesta sindacale!

c.c., che recita: “*salva diversa volontà delle parti, il debitore che nell’adempimento dell’obbligazione si avvale dell’opera di terzi, risponde anche dei fatti dolosi o colposi di costoro*”.

Dunque, il *dominus* principale, ovvero l’utilizzatore (o, a seconda dei casi, anche o esclusivamente il produttore o il manutentore), al pari del debitore, risponderebbe, salvo prova di una altrui responsabilità o corresponsabilità, della condotta del dispositivo intelligente per il solo fatto di essersi avvalso della sua opera, come avrebbe risposto qualora l’esecuzione della prestazione fosse stata affidata ad un uomo. Il tutto potrebbe essere compensato dal ricorso a polizze assicurative a tutela dell’utilizzatore (o del produttore o del manutentore).

Inoltre, il ricorso all’art. 2236 c.c., che prevede una limitazione di responsabilità nei soli casi di dolo o colpa grave per il prestatore di opera professionale, non può trovare un’applicazione generalizzata per ogni attività svolta con l’ausilio dei dispositivi intelligenti, poiché presupposto indefettibile per l’applicazione concreta di detto precetto codicistico, come è noto, è che la prestazione richiesta al professionista implichi la soluzione di problemi tecnici di particolare difficoltà e l’utilizzo di macchinari mossi da AI non necessariamente implica soluzioni di problemi tecnici di particolare difficoltà, ma, anzi, di mere questioni routinarie, banali o ripetitive.

La dottrina si è già interessata a questi complessi temi, mentre la giurisprudenza in tempi non lontani se ne occuperà facendo tesoro dei primi spunti scientifici in materia.

E in questi giudizi, fondamentale, se non dirimente, per acclarare la sussistenza di una responsabilità dell’utilizzatore di dispositivi intelligenti, sarà il grado di autonomia di questi ultimi ovvero la capacità degli stessi di compiere scelte ed attuarle nel mondo esterno indipendentemente da un controllo umano. Tanto maggiore, infatti, sarà il grado di autonomia della macchina, tanto minore sarà la responsabilità dell’utilizzatore (salvo ravvisare responsabilità del produttore o del progettista o del manutentore), essendo particolarmente difficile riscontrare il nesso di causalità tra l’operato di quest’ultimo e il danno cagionato al destinatario della scelta dell’AI (cittadino, paziente, parte di un giudizio, automobilista, viandante, ecc.). Poiché, peraltro, almeno allo stato, i dispositivi di intelligenza artificiale non sono così avanzati da poter demandare interamente agli stessi le decisioni inerenti il trattamento e l’esecuzione di attività umane (si pensi al campo delle cure mediche), sussisterà la responsabilità dell’utilizzatore ogniqualvolta il medesimo, fruendo dei dispositivi di AI, porrà in essere una condotta colposa ovvero quando la tecnologia intelligente abbia fornito indicazioni errate a causa dell’erroneo utilizzo della stessa da parte dell’uomo.

Tale previa e complessa allocazione della responsabilità per danni arrecati a terzi (ad es. utenti, clienti, pazienti, pedoni o automobilisti, ecc.) innesca altresì, vista la pluralità dei soggetti coinvolti nella realizzazione, alienazione, gestione, addestramento, uso della macchina dotata di intelligenza artificiale, problemi di rivalse interne fondate su rapporti contrattuali (ad es. legame tra produttore e utilizzatore o manutentore della macchina intelligente) o comunque sul reciproco affidamento nella bontà dei rispettivi contributi ideativi e lavorativi.

Né può trascurarsi un’ulteriore implicazione, in punto di responsabilità, legata, come evidenziato da autorevole dottrina (54), al non aver utilizzato i dispositivi intelligenti: si pensi alla condotta omissiva di sanitari che non abbiano fatto uso di sofisticate macchine intelligenti per la diagnosi e il trattamento dei propri pazienti. L’omesso utilizzo di dispositivi intelligenti avanzati e affidabili a fini diagnostici e terapeutici, potrebbe infatti essere foriera di una nuova forma di responsabilità per il medico (extracontrattuale) e per la struttura sanitaria (contrattuale, con successivi sviluppi di danno erariale innanzi alla Corte dei conti per danni a Ospedali pubblici risarciti in sede civile). Si pensi altresì al mancato utilizzo di programmi gestiti da algoritmi che possono prevedere terremoti o inondazioni foriere di morte.

Tutte tali tematiche vanno poi coordinate con i cennati profili assicurativi delle polizze sottoscritte dai vari soggetti sopra menzionati coinvolti in un danno cagionato a terzi, situazione che allargherà il grandangolo delle questioni giuridiche che verranno al pettine della magistratura. Il tema, oltre che nel campo sanitario (ove l’uso

(54) U. Ruffolo, *L’intelligenza artificiale in sanità*, cit., 503.

dell'AI è già molto diffuso sul piano diagnostico e interventistico) (55), è stato affrontato per le polizze per responsabilità civile da futura circolazione di veicoli guidati da AI (56).

In ogni caso, qualsiasi iniziativa giudiziaria da parte dei soggetti danneggiati presenta, oltre ai segnalati problemi di individuazione dei referenti normativi pertinenti e di ascrizione psicologica e causale della condotta dannosa al corretto autore, anche limiti probatori insiti nelle caratteristiche peculiari dell'IA, tra cui la sua complessità, parziale autonomia e ontologica opacità (il cosiddetto effetto “*scatola nera*”), che potrebbero rendere difficile o eccessivamente oneroso (attraverso costose Ct_u e Ct_i di parte) per i danneggiati identificare i responsabili e dimostrare la sussistenza dei requisiti necessari per un'azione basata sulla responsabilità extracontrattuale (più agevoli possono essere invece i binari della responsabilità contrattuale, ove ipotizzabile in capo a persone fisiche e giuridiche utilizzatrici di AI). Tali difficoltà potrebbero scoraggiare azioni risarcitorie: da qui l'intervento del Regolamento n. 1689 del 2024 in chiave preventiva *ex ante* sulla regolamentazione del rischio.

Per non demandare la soluzione di queste problematiche alla magistratura (quella reale e non quella artificiale!), va dunque in primo luogo e rapidamente ridotta l'incertezza normativa in materia attraverso l'adozione di una Direttiva in materia e vanno poi redatti pertinenti ed accurati contratti tra costruttore, venditore, utilizzatore e destinatario dei servizi resi dalle varie AI per regolare gli eventuali profili patologici. Ruolo decisivo avranno poi i prodotti assicurativi a tutela delle vittime, il cui premio (di complessa determinazione da parte degli attuari, in assenza di referenti ad oggi) varierà a seconda del grado di rischiosità di ogni singola AI.

Ecco perché va salutata con estremo favore la già segnalata adozione da parte del Parlamento europeo, su proposta della Commissione europea, per perseguire un approccio europeo coordinato volto a far fronte alla crescente diffusione dell'IA e delle tecnologie digitali, di un testo unitario, ancora *in itinere*, di una Direttiva sulla responsabilità per i danni causati dall'intelligenza artificiale (IA) (57) volta a facilitare sul piano europeo il risarcimento del danno a coloro che abbiano subito danni derivanti dall'impiego di sistemi di IA, al fine di accrescere la fiducia nell'IA e favorire gli investimenti nel mercato interno. Difatti, secondo un recente sondaggio, la responsabilità civile connessa all'utilizzo di sistemi di intelligenza artificiale è tra i primi tre ostacoli esterni all'utilizzo dell'IA da parte delle imprese europee nel loro complesso (33% del campione), mentre rappresenta il principale ostacolo per le imprese che non hanno ancora adottato sistemi di IA e hanno intenzione di farlo nei prossimi anni (43% del campione).

La direttiva IA *in itinere* è parte di un approccio europeo coordinato volto a far fronte alla crescente diffusione dell'IA e delle tecnologie digitali: l'adottato Regolamento sull'IA n. 1689 del 2024 (“Regolamento IA”), come detto, si concentra principalmente sul monitoraggio e sulla prevenzione dei danni, mentre la Direttiva IA mira ad

(55) Sull'uso dell'AI in sanità e sulle conseguenti responsabilità, oltre agli scritti menzionati in nt. 52, da ultimo v. V. Tenore (a cura di), *Studio sulla responsabilità disciplinare, civile, penale e amministrativo contabile del personale sanitario*, Roma, Anicia, 2023.

Oltre ai rilievi sviluppati sull'uso dell'AI in medicina nelle pregresse nt. 25 e 55, ulteriori esempi di applicazioni sono: dispositivi e servizi medici in aree di *screening* e diagnostica come l'omica e l'*imaging* medico, nuovi farmaci e vaccini, monitoraggio e trattamento delle persone, supporto alla cura del paziente (diagnosi e prognosi) e modelli predittivi delle esigenze sanitarie. Da segnalare il contenzioso vagliato da Tar Calabria, Catanzaro, Sez. II, 24 novembre 2011, n. 2092, su un sistema di “*badante digitale*”. Il dispositivo, secondo quanto illustrato in ricorso, ha il compito di creare un “*ambiente immersivo IOT*” – cioè internet delle cose – per monitorare, senza limitarle la libertà d'azione, i gesti di un malato affetto da limitazioni della sfera cognitiva, quale ad esempio il malato di Alzheimer o di altre demenze, riuscendo a prevenire e segnalare con anticipo gli atteggiamenti a rischio, potenziali fonti di danno fisico per il malato stesso, anche attraverso algoritmi di intelligenza artificiale. Il prototipo nella sua forma di MVP, *Minimun Viable Product*, è installato e funzionante presso la Rsa “Casa Amica” di Fossato Serralta di Catanzaro. E nel 2020 ha ottenuto il rilascio dell'attestato di brevetto quale invenzione industriale n. 102018000005262, “*Sistema e metodo di monitoraggio di pazienti con disturbi cognitivi*”.

(56) Cfr. la direttiva 2010/40/UE (anche nota come direttiva ITS) “Quadro generale per la diffusione dei sistemi di trasporto intelligenti nel settore del trasporto stradale e nelle interfacce con altri modi di trasporto”. Sul tema, v. A. Davola, R. Pardolesi, *In viaggio con il robot: verso nuovi orizzonti della r.c. auto (“driveless”)?*, in *Danno e resp.*, 2017, 625; E. Al Mureden, *Presentazione*, in G. Calabrese, *Il futuro del law and economics*, Milano, Giuffrè, 2018, XXXIV; U. Ruffolo, *Self driving car, auto drive-less e responsabilità*, in Id. (a cura di), *Intelligenza artificiale e responsabilità*, cit., 31.

(57) Già nel 2017, il Parlamento europeo, con la Risoluzione del 16 febbraio, affronta la materia della responsabilità civile nella robotica, attraverso le Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103 INL). Tendenzialmente, il Parlamento prevede di mantenere un approccio tradizionale, ma propone altresì di integrare le norme tradizionali attraverso delle novità, ossia l'introduzione di un sistema di registrazione dei robot, di una assicurazione obbligatoria e di un fondo di garanzia per i danni causati da robot non assicurati (punto 59, lett. a, b ed e), nonché di forme di responsabilità oggettiva (punti 53 e 54). Nel 2020, con il Libro bianco sull'IA (“Un approccio per l'eccellenza e la fiducia”), la Commissione europea propone di coniugare lo sviluppo e i rischi della nuova tecnologia con i valori europei. Il 28 settembre 2022 la Commissione europea ha adottato una proposta di direttiva sulla responsabilità per i danni causati dall'intelligenza artificiale (IA), al fine di accrescere la fiducia nell'IA e favorire gli investimenti nel mercato interno.

- Sul tema A. Bertolini, R. Limongelli, “*AI & Civil liability*”, in *Regulating Advanced Technologies: Policy Papers of the Jean Monnet Centre of Excellence on the Regulation of Robotics and AI*, Pisa, Il Campano, vol. 1, 2022, 39 ss.; E. Incampo, *Intelligenza artificiale e responsabilità: gli sviluppi comunitari*, in <www.diritto.it>, 3 giugno 2024. Sul rapporto tra regolazione del rischio *ex ante* (operata dal Regolamento n. 1689 del 2024) e risarcimento del danno *ex post*, v. G. Smorto, *Distribuzione del rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale*, cit., 217.

armonizzare il regime di responsabilità applicabile nei casi in cui i sistemi di IA causino danni, siano essi sistemi ad alto o a basso rischio. Secondo la Commissione, infatti, le norme nazionali vigenti in materia di responsabilità, in particolare per colpa, non sono adatte a gestire le azioni di responsabilità per danni causati da prodotti e servizi basati sull'IA.

In tempi recenti, seppur per un peculiare profilo risarcitorio, la Corte di cassazione ha vagliato una controversia relativa alla responsabilità per danno causato da un sistema di intelligenza artificiale di *rating* reputazionale per illecito trattamento di dati personali. In questo caso, la causa del danno da illecito trattamento di dati personali da parte di un sistema di intelligenza artificiale è stata individuata nella assenza di trasparenza circa l'algoritmo utilizzato dal sistema stesso per determinare il *rating*.

Ma altre questioni sono sorte o potrebbero nascere, sul piano civilistico o comunque innanzi all'autorità giudiziaria ordinaria, in ordine: a) a contenziosi su sanzioni inflitte dal Garante della *privacy* per trattamento di dati personali senza consenso dell'interessato acquisiti da una piattaforma *web* tramite una AI (58); b) alla brevettabilità di opere di ingegno o invenzioni frutto di AI, tematica affrontata in alcuni contenziosi in altri Paesi (59) e dalla stessa Cassazione in materia di tutela del diritto d'autore per opere d'ingegno elaborate tramite AI (60); c) a sinistri stradali o aerei legati ad incidenti tra veicoli guidati da AI; d) all'utilizzo di AI per decisioni più obiettive, uniformi e meno "umoralì" per la determinazione dell'assegno di mantenimento in caso di separazioni e divorzi, valutando i dati istruttori in atti.

Ma l'evoluzione gestionale del personale (selezione, assegnazione mansioni, trasferimenti, conferimento incarichi, ristrutturazioni, procedimenti disciplinari ecc.) tramite AI, sia in contesti pubblici che privati, pone anche nuove *questioni lavoristiche* (61): ad esempio la scarsa trasparenza dei sistemi algoritmici nelle scelte

(58) Il tema è stato affrontato da Cass., Sez. I, 25 maggio 2021, n. 14381, secondo cui in tema di trattamento di dati personali, il consenso è validamente prestato solo se espresso liberamente e specificamente in riferimento ad un trattamento chiaramente individuato; ne consegue che nel caso di una piattaforma *web* (con annesso archivio informatico) preordinata all'elaborazione di profili reputazionali di singole persone fisiche o giuridiche, incentrata su un sistema di calcolo con alla base un algoritmo finalizzato a stabilire punteggi di affidabilità, il requisito della consapevolezza non può considerarsi soddisfatto ove lo schema esecutivo dell'algoritmo e gli elementi di cui si compone restino ignoti o non conoscibili da parte degli interessati. Sul rapporto tra AI e *privacy*, v. G. Natale, *Artificial Intelligence and privacy right*, cit.

(59) Il riferimento è alla decisione della Corte federale australiana nel caso *Thaler vs. Commissioner of Patents*, che ha negato la possibilità di brevettare un'invenzione creata da un sistema di intelligenza artificiale perché privo della personalità giuridica, ossia della capacità di essere titolare di situazioni giuridiche soggettive.

Si è poi già ricordato nel par. 1 che nei confronti di *Open AI*, creatrice di Chat-GPT, è stata però di recente intentata in un tribunale federale californiano, a luglio 2023, dallo studio californiano Clarkson, una *class action* per asserita violazione del diritto d'autore e della *privacy* di milioni di cittadini, in quanto il sistema di intelligenza artificiale "generativa" fa allenare i suoi algoritmi con masse di dati e informazioni acquisibili in fonti aperte *online*, ovvero quelle consultabili da chiunque. Altro contenzioso statunitense riguarda l'azione intrapresa da importante testata giornalistica per l'uso di propri articoli ripresi da ChatGPT.

Evidente è dunque l'ampia casistica di novelli contenziosi civili originati dall'uso di AI.

(60) Cass., Sez. I, 16 gennaio 2023, n. 1107, in *Guida dir.*, 2023, 10, relativa alla riproduzione non autorizzata dell'immagine creativa di un fiore che può costituire violazione dei diritti d'autore di chi ha realizzato l'immagine, anche nell'ipotesi in cui si sia fatto ricorso nel processo creativo all'utilizzo di un *software*. La Suprema Corte, sposando una tesi già espressa in più sedi e conferendole autorevolezza, ha precisato incidentalmente che il ricorso alla tecnologia digitale per la realizzazione di un'opera non preclude di per sé la possibilità di qualificare l'opera come opera dell'ingegno, a meno che, all'esito di un accertamento di fatto in cui il tasso di creatività sia stato scrutinato con rigore, non risulti che l'utilizzo della tecnologia abbia assorbito l'elaborazione creativa dell'artista. Con tale precisazione la Cassazione contribuisce in maniera significativa al dibattito in essere da qualche anno circa la proteggibilità ai sensi del diritto d'autore delle opere generate da sistemi di intelligenza artificiale: sulla scorta del criterio indicato, occorrerà distinguere tra opere generate attraverso sistemi di intelligenza artificiale rispetto alle quali l'elaborazione creativa dell'uomo assuma un rilievo significativo e opere generate da sistemi di intelligenza artificiale rispetto alle quali l'apporto creativo dell'uomo sia marginale. Solo nel primo caso, infatti, l'opera sarà tutelabile secondo le categorie tradizionali del diritto d'autore, e tanto il diritto morale di essere riconosciuto autore, quanto i diritti di sfruttamento economico dell'opera andranno riconosciuti in capo all'artista persona fisica.

In ogni caso ha chiarito l'Ufficio europeo brevetti, con decisione 27 gennaio 2020, cit., che i sistemi o le macchine di AI non sono titolari di diritti, anche ai fini della procedura di brevettazione di una invenzione, perché non hanno una personalità giuridica paragonabile alle persone fisiche o giuridiche.

(61) Sulle diverse tematiche lavoristiche derivanti dall'uso datoriale di AI, v. A. Lo Faro, *Algorithmic Decision Making e gestione dei rapporti di lavoro: cosa abbiamo imparato dalle piattaforme*, in <www.federalismi.it>, 5 ottobre 2022; P. Loi, *Il rischio proporzionato nella proposta di regolamento sull'IA e i suoi effetti nel rapporto di lavoro*, ivi, 8 febbraio 2023; M. Barbera, *"La nave deve navigare". Rischio e responsabilità al tempo dell'impresa digitale*, in *Labour & Law Issues*, 9(2), 2023, 1 ss.; L. Tebano, *Poteri datoriali e dati biometrici nel contesto dell'AI Act*, in <www.federalismi.it>, 18 ottobre 2023; A. Donini, *Informazione sui sistemi decisionali e di monitoraggio automatizzati: poteri datoriali e assetti organizzativi*, in *Diritti lavori mercati*, 2023, 85; S. Ciucciovino, *La disciplina nazionale sulla utilizzazione della intelligenza artificiale nel rapporto di lavoro*, in <www.lavorodirittieuropa>, 29 gennaio 2024; M. Biasi (a cura di), *Diritto del lavoro e intelligenza artificiale*, Milano, Giuffrè, 2024, 147 ss.; M. Peruzzi, *L'impatto dell'IA nella selezione del personale, negli annunci di lavoro mirati a filtrare le candidature, nella determinazione della "reputazione" della persona che lavora e nell'assegnazione dei compiti*, in <www.lavorodirittieuropa>, 25 luglio 2024.

organizzative datoriali ha originato nuovi conflitti per condotta antisindacale (62). Si aprono inoltre scenari, sempre lavoristici, di nuove forme di discriminazione “algoritmiche” e di spersonalizzazione delle scelte datoriali soprattutto in settori ove la valutazione umana di circostanze fattuali è basilare (ad esempio in materia di responsabilità disciplinare).

Le riflessioni sopra svolte hanno riguardato i profili civili risarcitori legati all’uso delle AI. Ma anche sotto il profilo della *responsabilità penale* si pongono delicati problemi probatori e di ascrizione dei possibili reati frutto di utilizzo di intelligenze artificiali. Si pensi solo alla rilevanza della distinzione tra dolo eventuale e colpa cosciente in capo ad un utilizzatore o programmatore di AI.

A fronte del principio della irresponsabilità penale di una mera *res* e del consequenziale canone *machina delinquere non potest*, il giurista israeliano Gabriel Hallevy ha ipotizzato tre astratti paradigmi di responsabilità penale applicabili a un’entità dotata di intelligenza artificiale: a) *Perpetration through another*, ovvero il programmatore o l’utente finale potrebbero essere ritenuti responsabili di aver istruito direttamente l’entità IA a commettere il crimine. La responsabilità dolosa è imputabile esclusivamente all’essere umano che ha agito attraverso un’entità dotata di apprendimento automatico o per mezzo di un algoritmo preimpostato; b) *Natural probable consequence*, ovvero il programmatore o l’utente finale non hanno pianificato la commissione di alcun crimine. Secondo lo schema della colpa cosciente, il reato commesso materialmente dalla macchina costituiva una conseguenza probabile, non presa in considerazione dall’uomo che l’ha progettata; c) *Direct liability*, ovvero il programmatore o l’utente finale non hanno nessuna responsabilità, che ricade esclusivamente in capo all’entità IA.

Solo il primo paradigma, *perpetration through another*, prevede una responsabilità penale unicamente umana; mentre il secondo e il terzo paradigma presuppongono il riconoscimento di una personalità giuridica con una conseguente responsabilità penale delle entità IA. Il giurista Havelly ipotizza un parallelismo tra la responsabilità penale degli enti, introdotta nell’ordinamento italiano con il d.lgs. 8 giugno 2001, n. 231, che ha superato il brocardo *societas delinquere non potest*, e l’introduzione della responsabilità penale dell’entità IA: gli enti rispondono dei reati posti in essere nell’interesse dell’ente da chi svolge funzioni di rappresentanza, di amministrazione, di direzione dello stesso. Diversamente, altri ritengono che l’entità IA non può essere concepita come centro autonomo di imputazione, né come entità attraverso la quale si manifestano volontà umane.

Un cenno doveroso va poi fatto alle nuove problematiche che nel *diritto tributario* porranno, sul piano sostanziale e processuale, accertamenti fiscali fondati su algoritmi, che dovranno ovviamente palesare i criteri seguiti a tutela del diritto alla difesa del contribuente.

Da ultimo, l’uso di AI in *fase processuale* (istruttoria e decidente), come meglio si vedrà nel conclusivo par. 4, al quale è sufficiente rinviare.

3.1. Segue: *Contenziosi amministrativi e contabili nascenti dall’uso di intelligenza artificiale*

Non trascurabili sono poi i contenziosi amministrativi innanzi al g.a., destinati a crescere, nascenti da uso di AI da parte delle pubbliche amministrazioni e ai possibili risvolti di responsabilità amministrativo-contabile derivante da danni erariali cagionati dall’uso di AI da parte di enti pubblici.

Premesso che l’utilizzo di una procedura informatica che conduca direttamente alla decisione finale della pubblica amministrazione, pur presentando rischi (63), dev’essere incoraggiato, comportando numerosi vantaggi

(62) Ad esempio, Trib. Palermo, Sez. lav., 20 giugno 2023 (sul tema v. anche Trib. Palermo, ord. 3 aprile 2023, n. 14491 e Trib. Bologna 31 dicembre 2020) ha vagliato casistiche sulla trasparenza dei sistemi algoritmici come nuova frontiera del conflitto sindacale, acclarando la natura antisindacale della violazione dell’obbligo di informazione. Le organizzazioni sindacali, nella fattispecie *sub iudice*, dopo aver formalmente richiesto alla società di consegne a domicilio di comunicare le informazioni circa la logica algoritmica dei sistemi decisionali e di monitoraggio automatizzati, utilizzati nell’organizzazione del lavoro dei rider, lamentavano la natura antisindacale della condotta della società, che dapprima aveva ignorato la richiesta di adempimento dell’obbligo informativo e in seguito, dopo la notifica del ricorso *ex art. 28* Statuto dei lavoratori, aveva comunicato alle organizzazioni sindacali solo alcune informazioni, ritenute non soddisfatti dalle ricorrenti. Respingendo le eccezioni e le obiezioni della resistente, il Tribunale di Palermo ha dichiarato antisindacale la condotta in questione, ordinando alla società di informare le rappresentanze sindacali circa scopi, finalità, logica e funzionamento dei sistemi automatizzati in uso per l’organizzazione del lavoro. A presidio della pronuncia, inoltre, il giudice ha previsto il versamento da parte dell’impresa alle organizzazioni sindacali di 200,00 euro giornalieri, a partire dal quindicesimo giorno dalla comunicazione del decreto, in caso di inadempimento del prescritto obbligo di informazione.

(63) Cons. Stato, Sez. atti normativi, 12 gennaio 2023, n. 137, ben rimarca: “La Sezione ritiene, invece, che è necessaria un’attenta valutazione del nutrito e complesso contenzioso insorto in relazione alla digitalizzazione sia delle modalità di accesso alle prove concorsuali che del loro svolgimento. A mero titolo esemplificativo, e senza alcuna pretesa di esaustività, si richiamano le molteplici pronunce del giudice amministrativo di primo e di secondo grado in materia di esclusioni dalle procedure concorsuali a causa di errori e/o omissioni nella compilazione della domanda telematica e dei connessi principi di affidamento digitale e di soccorso istruttorio, elaborati dalla giurisprudenza (cfr. Cons. Stato, Sez. VI, 15 marzo 2021, n. 2226; Tar Lazio, Sez. II, 19 aprile 2022, n. 4664), così come quelle concernenti i casi di crash di rete ovvero di altre tipologie di difficoltà telematiche verificatesi nel corso di procedure concorsuali da remoto, quando siano riconducibili a un malfunzionamento della connessione internet o a ragioni tecniche e non a negligenza o a malafede del partecipante (cfr. Cons. Stato, Sez. VI, 1 luglio 2021, n. 5008) o ancora quelle afferenti all’incidenza della digitalizzazione sulla valutazione della

quali, ad esempio, la notevole riduzione della tempistica procedimentale per operazioni ripetitive, la prevedibilità maggiore degli esiti rimarcata da tempo da attenta dottrina (64) (le stesse finalità in punto di “prevedibilità”, come si è visto, sono alla base degli indirizzi favorevoli alla giustizia predittiva tramite uso di AI), l’esclusione di interferenze dovute a negligenza del funzionario o a conflitti di interesse non palesati (nonostante l’obbligo sancito dal d.p.r. n. 62/2013) e la conseguente maggiore garanzia di imparzialità, sono sorti diversi contenziosi derivanti dal sempre più frequente ricorso a procedimenti amministrativi telematici guidati da AI.

La possibilità di delegare ad una macchina l’elaborazione di atti vincolati (rilascio di certificati, duplicati di documenti smarriti, pagamenti vari, iscrizioni a scuole e università, ecc.) è questione pacifica e difficilmente foriera di contenziosi. È invece sulle scelte discrezionali della p.a. che insorgono problemi, *in primis* sulla correttezza o meno del ragionamento seguito dall’algoritmo per pervenire alla decisione, ascritta soggettivamente ad un funzionario, ma assunta da una macchina (non essendo ipotizzabile una “volontà della macchina”) (65).

Oltre a vagliare frequenti giudizi derivanti da esclusioni di domande di concorso, di gara o di finanziamenti pubblici presentate in via telematica e illegittimamente respinte dal gestore-AI con risvolti anche risarcitori (66), il Consiglio di Stato ha poi chiarito che la regola tecnica che governa ciascun algoritmo resta pur sempre una

prova concorsuale. Merita anche di essere ricordato che la giurisprudenza è costante nell'affermare che l'utilizzo delle tecnologie informatiche nella procedura amministrativa non può pregiudicare i partecipanti, dovendo la p.a. farsi carico di eventuali disfunzioni del sistema. A tale riguardo la Sezione osserva che l'analisi della giurisprudenza formatasi sulla materia in esame consentirebbe anche di individuare misure idonee a sterilizzare i rischi di contenzioso attraverso la previsione di presidi idonei di verifica e controllo delle tecnologie digitali, necessariamente diversificati in relazione alla tipologia di criticità da fronteggiare: criticità che possono scaturire dall'uso di tecnologie di intelligenza artificiale ovvero discendere da asserite problematiche connesse al funzionamento delle piattaforme e alla loro ispezionabilità. Al fine di perseguire e valorizzare le potenzialità insite nel ricorso alle nuove tecnologie, in termini di velocizzazione e di semplificazione delle modalità di presentazione delle domande di partecipazione e anche dello svolgimento delle stesse prove selettive, la Sezione ritiene che sia necessaria e utile una panoramica dei principali profili di rischio derivanti dalla loro utilizzazione, come stigmatizzati nelle pronunce del giudice amministrativo”.

Per un’accurata analisi dell’evoluzione dell’uso dell’AI da parte della p.a., v. L. Viola, *L’intelligenza artificiale nel procedimento e nel processo amministrativo*, cit. Sulla trasformazione delle professioni legali (avvocati, notai, giudici, professori) a seguito della rivoluzione digitale, v. L. Cappello et al., *La trasformazione delle professioni legali. Legal engineering, blockchain, metaverso, IoT e altre tecnologie nella digital economy*, Torino, Giappichelli, 2023.

(64) Un lungimirante studio di A. Police, *La predeterminazione delle decisioni amministrative. Gradualità e trasparenza nell’esercizio del potere discrezionale*, Napoli, Esi, 1997, 247 ss., offre importanti spunti in merito alla predeterminazione dell’azione pubblica che accrescerebbe anche il grado di prevedibilità delle decisioni amministrative, facendo sorgere, di conseguenza, in capo agli amministratori un legittimo affidamento in ordine all’esito dell’azione stessa.

(65) Come ben rimarca L. Viola, *L’intelligenza artificiale nel procedimento e nel processo amministrativo*, cit., 27, per favorire la gestione automatizzata dell’agere della p.a., a prescindere sia dalla difficoltà nel definire casi facili (automatizzabili) e casi difficili (da devolvere ad uomini) nell’attività amministrativa, sia dalla fluidità e mutevolezza nel tempo nella nozione di atto vincolato, attraverso la teoria dell’autolimita, la p.a. «potrebbe avere la possibilità non solo di ridurre il tasso di discrezionalità delle sue valutazioni finali (circostanza che appare assai rilevante, ai fini dell’automazione dell’emanazione del provvedimento finale), ma anche di trasformare valutazioni originariamente caratterizzate dal carattere discrezionale in vincolate, per effetto proprio della predeterminazione di criteri generali di valutazione caratterizzati da un carattere sufficientemente univoco da essere suscettibile di automazione. In buona sostanza, la natura discrezionale di un potere di valutazione non costituisce pertanto ostacolo all’automazione, ove sia possibile o obbligatorio (in applicazione delle previsioni in materia di autolimita della p.a.) prevedere dei criteri generali di valutazione che vengano a “ridurre” il tasso di discrezionalità fino ad un livello compatibile con l’automazione». Lo studio di Viola analizza poi le tesi favorevoli e radicalmente ostative all’uso di AI nella redazione di atti amministrativi e giunge ad una tesi favorevole ad una ragionevole simbiosi tra uomo e sistema informatico finalizzata ad affidare all’elaborazione automatica particolari fasi o momenti delle attività giuridiche che si affiancano e integrano con quelli affidati all’uomo, per migliorare l’efficienza della p.a. o anche per correggere, con programmi-sentinella, errori o incongruenze di scelte fatte dall’uomo (si pensi ad alcuni programmi di revisione di secondo e terzo grado utilizzati dall’Inps o dalla Agenzia delle entrate per monitorare il lavoro e le scelte discrezionali o valutative dei propri funzionari nel concedere benefici non spettanti).

(66) V. il caso vagliato da Cons. Stato, Sez. III, 28 dicembre 2020, n. 8435, relativo dalla illegittima esclusione del Codacons da una procedura telematica di partecipazione ad un finanziamento pubblico per mancata sottoscrizione di un allegato alla domanda (ma la cui paternità, riconducibile al Codacons, era palese per la procedura telematica seguita), senza applicazione del soccorso istruttorio e foriera di un danno patrimoniale. Afferma il g.a.: “Si realizza così, in definitiva, di una moderna forma di partecipazione al procedimento amministrativo telematico, così evolutosi in modalità automatizzata, talora per effetto di un software di intelligenza artificiale, che interagisce col partecipante-utente senza l’intervento di alcun operatore o persona fisica. Se dunque si applicano questi principi al caso di specie, può riconoscersi come la procedura telematica in questione contenga sicuri elementi da cui desumere esattamente la riconducibilità all’autore della dichiarazione-attestazione ritenuta irregolare o incompleta, e ciò a prescindere dall’evento formale dell’avvenuta sottoscrizione materiale della dichiarazione. In particolare, ciò si ricava da quanto previsto dal disciplinare telematico, secondo cui la partecipazione alla procedura di gara, e la presentazione della relativa domanda, erano possibili solo attraverso l’accesso al sistema effettuato mediante la creazione di un apposito account identificativo, del resto assorbito, quanto a certezza sulla paternità della domanda e degli atti ad essa allegati, dalla provenienza telematica, e dunque non certamente anonima, degli atti trasmessi tramite upload. Il caricamento nella piattaforma della domanda e dei suoi allegati, documentabili ex se attraverso i software di trasmissione e ricezione documentale, rendevano quindi certa l’identità del presentatore e dunque la provenienza di ogni atto pervenuto in tale modalità”.

regola amministrativa generale, costruita dall'uomo e non dalla macchina, per poi essere *in concreto* applicata da quest'ultima. Il giudice amministrativo ha pertanto più volte ribadito (67) che il ricorso alla funzione algoritmica all'interno del procedimento amministrativo non è vietato di per sé, neppure in relazione ai procedimenti caratterizzati da discrezionalità, anche tecnica, a condizione che si rispettino determinati requisiti, derivanti sia dai principi di diritto interno che dalle norme del diritto europeo (68): si pensi a procedimenti di gara, di concorso, di trasferimento, di erogazione di finanziamenti, di assegnazione di alloggi o di sedi di servizio. Ma non sono mancati, soprattutto nelle complesse procedure di assegnazione dei docenti del Ministero dell'istruzione, indirizzi più cauti della giurisprudenza (69).

Tuttavia, il ricorso all'algoritmo in funzione istruttoria, integrativa e servente della decisione umana, ovvero anche in funzione parzialmente decisionale nei procedimenti a basso tasso di discrezionalità, non può mai comportare un abbassamento del livello delle tutele garantite dalla legge sul procedimento amministrativo, e in particolare di quelle sulla individuazione del responsabile del procedimento, sull'obbligo di motivazione (70), sulle garanzie partecipative e sulla c.d. non esclusività della decisione algoritmica. Tra le indicate garanzie assume primaria importanza il rispetto del principio di trasparenza, che, com'è noto, trova un immediato corollario

(67) Cons. Stato, Sez. VI, 8 aprile 2019, n. 2270, in *Dir. amm.*, 2019, 784, con nota di R. Ferrara, *Il giudice amministrativo e gli algoritmi, note estemporanee a margine di un recente dibattito giurisprudenziale*, annotata da G. Pesce, *Il Consiglio di Stato ed il vizio dell'opacità dell'algoritmo tra diritto interno e diritto sovranazionale*, in <www.giustizia-amministrativa.it>, 16 gennaio 2020; L. Musselli, *La decisione amministrativa nell'età degli algoritmi: primi spunti*, in <www.medialaws.eu>, 16 marzo 2020; v. anche Cons. Stato 13 dicembre 2019, nn. 8472, 8473, 8474, e, da ultimo, Cons. Stato, Sez. VI, 28 aprile 2023, n. 4297, secondo cui “il ricorso a strumenti informatici nelle procedure amministrative costituisce una modalità agevolata di istruttoria, senza che il singolo strumento – per quanto qualificabile in termini di intelligenza artificiale – possa, da un lato, derogare alle regole normative ed ai criteri posti a presupposto della singola procedura e, dall'altro lato, essere sottratto alla trasparenza nonché alla imputabilità all'amministrazione procedente. Quest'ultima è chiamata a verificare la correttezza del funzionamento dello strumento istruttorio utilizzato e la relativa coerenza agli obiettivi ed alle regole dettate, in coerenza al principio di legalità, per l'esercizio del potere in esame”.

Tale indirizzo del g.a. viene ripreso anche da Trib. Bari, Sez. lav., 30 maggio 2023, n. 1565, in un contenzioso su scavalchi in graduatorie di docenti, e da Trib. Trani, Sez. lav., 30 gennaio 2023, n. 2210 (sul medesimo tema dei trasferimenti dei docenti), secondo cui «non può presumersi che dalla circostanza di aver “affidato” la combinazione dei dati della procedura di mobilità a un sistema di intelligenza artificiale derivi, per ciò solo, la legittimità della procedura, occorrendo dimostrare che tale sistema abbia effettivamente funzionato in base a parametri che siano in definitiva rispettosi dei principi generali dell'azione amministrativa. L'illegittimità del ricorso all'algoritmo determina quindi l'illegittimità dell'intera procedura di mobilità straordinaria proprio in quanto risulta che tale procedura è stata basata sull'uso dell'algoritmo senza provare che ciò, al contempo, assicurasse la legittimità della scelta nei trasferimenti ed escludesse soluzioni illogiche e irrazionali».

Sul tema, in dottrina, v.: D.U. Galetta, J.G. Corvalán, *Intelligenza Artificiale per una pubblica amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in <www.federalismi.it>, 6 febbraio 2019; G. Orsoni, E. D'Orlando, *Nuove prospettive dell'amministrazione digitale: Open data e algoritmi*, in *Istituzioni del federalismo*, 2019, 593; E. Carloni, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, in *Riv. dir. amm.*, 2020, 273; F.F. Pagano, *Pubblica amministrazione e Innovazione tecnologica*, in *La Rivista Gruppo Pisa*, 2021, 3, 84; S. Vernile, *Verso la decisione amministrativa algoritmica*, in <www.medialaws.eu>, 15 luglio 2020; L. Viola, *L'intelligenza artificiale nel procedimento e nel processo amministrativo*, cit.; A. Masucci, *Procedimento amministrativo e nuove tecnologie*, cit., 93-94; C. Romeo, *L'avatar, il metaverso e le nuove frontiere del lavoro: traguardo o recessione*, in *Lavoro giur.*, 2023, 471 ss.; S. Foà, *Intelligenza artificiale e cultura della trasparenza amministrativa. Dalle “scatole nere” alla “casa di vetro”*?, in *Dir. amm.*, 2023, 515 ss.

(68) Cons. Stato, Sez. VI, 4 febbraio 2020, n. 881; n. 2270/2019, cit.; nn. 8472, 8473 e 8474/2019, cit.; 8 settembre 2022, n. 6236; Sez. III, 25 novembre 2021, n. 7891. Afferma Cons. Stato, Sez. V, 20 ottobre 2008, n. 5114, che “il ricorso a procedure automatizzate, ai fini del reclutamento dei pubblici impiegati, è stato ritenuto pienamente legittimo dalla giurisprudenza amministrativa e costituisce ormai pratica ampiamente diffusa”.

Ha ben chiarito Tar Lazio, Roma, Sez. III-bis, 22 marzo 2017, n. 3769, che la scelta operata dall'amministrazione di ricorrere a strumenti elettronici di conduzione di procedimenti fosse qualificabile alla stregua di una manifestazione di discrezionalità che, in quanto tale, sarebbe stata sindacabile solo in caso di illogicità manifesta, ravvisabile nell'ipotesi in cui gli strumenti adoperati fossero inadeguati, giacché obsoleti, mal realizzati, o perché le peculiari funzioni esercitate fossero di complessità tale da non poter essere demandate ad una formula.

(69) Tar Lazio, Roma, Sez. III-bis, 10 settembre 2018, n. 8224, giudicando sugli esiti di un procedimento di mobilità condotto dal Miur, ha ritenuto illegittimo che si fosse “demandato ad un impersonale algoritmo lo svolgimento dell'intera procedura di assegnazione dei docenti”, ed ha altresì affermato che la eventuale complessità dell'azione svolta non può comunque costituire una valida giustificazione per giungere alla sua “devoluzione ad un meccanismo informatico o matematico del tutto impersonale e orfano di capacità valutazionali delle singole fattispecie concrete, tipiche invece della tradizionale e garantistica istruttoria procedimentale che deve informare l'attività amministrativa”. Tale indirizzo del g.a. viene ripreso anche da Trib. Bari, Sez. lav., n. 1565/2023, cit., in un contenzioso su scavalchi in graduatorie di docenti, e da Trib. Trani, Sez. lav., n. 2210/2023, cit. (sul medesimo tema dei trasferimenti dei docenti).

(70) La dottrina favorevole all'obbligo di motivare gli atti amministrativi ad elaborazione elettronica cita spesso Cons. Stato, Sez. VI, 24 ottobre 2000, n. 5682, che ha certo affermato la sindacabilità per difetto di motivazione dell'attività amministrativa in forma elettronica (in quel caso, si trattava degli elaborati di un concorso), ma solo come *obiter dictum*. Sul tema delle garanzie della l. n. 241 a procedimenti elaborati tramite AI, v. A. Masucci, *L'atto amministrativo informatico. Primi lineamenti di una ricostruzione*, Napoli, Jovene, 1993, 99; L. Viola, *L'intelligenza artificiale nel procedimento e nel processo amministrativo*, cit.; A. Masucci, *Procedimento amministrativo e nuove tecnologie*, cit.

nell'obbligo di motivazione degli atti amministrativi ex art. 3 l. n. 241/1990, e che non può essere soppresso né ridotto solo per la presenza di un algoritmo all'interno dell'iter procedimentale.

Secondo il giudice amministrativo, il fatto che il provvedimento venga emanato sulla scorta di una complessa operazione di calcolo produce l'effetto di rafforzare, per certi versi, l'obbligo motivazionale in capo all'amministrazione, la quale dovrà rendere la propria decisione finale non solo conoscibile, ma anche comprensibile. Occorre, secondo il g.a., spostare l'attenzione a monte, sulla costruzione dell'algoritmo; su come i parametri dell'algoritmo vengono scelti (operazione di per sé soggettiva) e come si combinano tra loro; e ancor prima su come i termini assunti quale parametro siano stati realizzati. La questione dell'individuazione dei termini da assumersi per la costruzione dell'algoritmo indica il momento in cui si opera la scelta caratterizzata da discrezionalità, sì che a queste fasi preliminari alla nascita dell'algoritmo devono essere anticipate le garanzie che devono accompagnare ogni scelta dell'amministrazione.

Pertanto, ad avviso del giudice amministrativo, fondamentale è a tal fine la garanzia di trasparenza, volta ad assicurare la conoscibilità della costruzione dell'algoritmo, anche, eventualmente, in funzione del sindacato sull'atto adottato sulla base dello stesso: *“la decisione amministrativa automatizzata impone al giudice di valutare in primo luogo la correttezza del processo informatico in tutte le sue componenti: dalla sua costruzione, all'inserimento dei dati, alla loro validità, alla loro gestione”* (71). In caso di decisione fondata su algoritmo, si richiede pertanto che sia assicurata una *“declinazione rafforzata del principio di trasparenza”*, intesa come *“piena conoscibilità della regola espressa in un linguaggio differente da quello giuridico”*. La *“caratterizzazione multidisciplinare”* dell'algoritmo comporta, secondo il corretto ragionamento del g.a., che la sua comprensione non richiede solo competenze giuridiche, ma anche tecniche, informatiche, statistiche, amministrative, sicché non esime dalla necessità che la *“formula tecnica”*, che di fatto rappresenta l'algoritmo, sia corredata da spiegazioni che la traducano nella *“regola giuridica”* ad essa sottesa e che la rendano leggibile e, quindi non solo *“conoscibile”*, ma anche *“comprensibile”*.

Non va poi dimenticato il principio di non esclusività della decisione algoritmica in base all'art. 22 del Regolamento generale sulla protezione dei dati, secondo cui *“L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona”*.

Da tale referente normativo si rinviene la conferma del ruolo ancillare dell'AI rispetto all'uomo: il programma, come ben colto da attenta dottrina (72), dovrebbe essere inquadrato come una sorta di responsabile (dell'istruttoria) del procedimento, che presenta una bozza di atto all'organo-uomo titolare del potere di svolgimento della funzione. Il programma rappresenterebbe una sorta di *“funzionario virtuale”* un *“travet-robot”*, che opera in funzione strumentale rispetto ad un dirigente chiamato a verificare se la macchina abbia operato bene o commesso errori.

Il Consiglio di Stato ha poi *ex professo* operato una chiara distinzione concettuale tra algoritmo e intelligenza artificiale in un interessante contenzioso volto ad ottenere la fornitura di un dispositivo medico (*pacemaker*) con elevato grado di automazione (73).

(71) I passi virgolettati sono tratti da Cons. Stato, Sez. VI, n. 2270/2019, cit.

Da ultimo Tar Campania, Napoli, 14 novembre 2022, n. 7703 (in <www.amministrativamente.com>, 2023, con nota di V. Dardano, *I limiti nell'utilizzo delle decisioni automatizzate e la legalità algoritmica*, annotato da C. Elia, *Sull'uso degli algoritmi nell'ambito del procedimento amministrativo*, in *Gazzetta forense*, 2023, 1, 198) ha chiarito, nell'annullare un provvedimento dell'Agea di recupero somme indebitamente versate ad un agricoltore in misura eccessiva e ricalcolate con un algoritmo, come *«il rispetto del principio di trasparenza imponga un indefettibile obbligo motivazione a carico della pubblica amministrazione, che si declina nella conoscibilità e nella comprensibilità del meccanismo algoritmico utilizzato. E ciò al fine di consentire, da un lato, il pieno esercizio del diritto di difesa da parte del soggetto inciso dal provvedimento, ai sensi degli artt. 24 e 113 Cost., dall'altro, il pieno sindacato di legittimità da parte del giudice amministrativo. Altro principio affermato in sede europea è quello di non esclusività della decisione algoritmica (art. 22 Gdpr), il quale attribuisce al destinatario degli effetti giuridici di una decisione automatizzata il diritto a che tale decisione non sia basata unicamente sul processo automatizzato, affidando al funzionario responsabile il compito di controllare, e quindi validare o, al contrario, smentire la decisione automatica. Ne deriva che, come sopra rilevato, la decisione adottata con ricorso all'algoritmo vede sempre la necessità che sia l'amministrazione a compiere un ruolo ex ante di selezione e verifica, anche per mezzo di costanti test, aggiornamenti e modalità di perfezionamento dell'algoritmo. Ma soprattutto, in tali casi si deve contemplare la possibilità che sia il giudice a “dover svolgere, per la prima volta sul piano ‘umano’, valutazioni e accertamenti fatti direttamente in via automatica”, con la conseguenza che la decisione robotizzata “imponga al giudice di valutare la correttezza del processo automatizzato in tutte le sue componenti (Cons. Stato, Sez. VI, sent. 2270/2019)»*.

(72) V. Dardano, *op. cit.*

(73) Chiarisce Cons. Stato, Sez. III, n. 7891/2021, cit., che l'algoritmo è una sequenza finita di istruzioni, ben definite e non ambigue, così da poter essere eseguite meccanicamente e tali da produrre un determinato risultato. Tale nozione, quando è applicata a sistemi tecnologici, è ineludibilmente collegata al concetto di automazione, ossia a sistemi di azione e controllo idonei a ridurre l'intervento umano, che dipendono dalla complessità e dall'accuratezza dell'algoritmo che la macchina è chiamata a processare. Cosa diversa, invece, è l'intelligenza artificiale, dal momento che, in questo caso, l'algoritmo contempla meccanismi di *machine learning* e crea un sistema che non si limita solo ad applicare le regole *software* e i parametri

Ma il complesso Tar-CdS ha anche affrontato in senso positivo il rilevante e connesso tema della concezione “attizio-provvedimentale” o meno del *software*, richiamando l’attenzione sul dato per cui è con il *software* che la pubblica amministrazione costituisce, modifica ed estingue situazioni giuridicamente rilevanti, mentre resta del tutto irrilevante che l’algoritmo sia compilato in un linguaggio di programmazione incomprensibile al funzionario e al cittadino interessati (74).

Ben potrebbe poi il giudice amministrativo doversi interessare di danni da provvedimento illegittimo elaborato tramite intelligenza artificiale, in applicazione del noto brocardo *causa causae est causa causati*.

La tematica è stata già affrontata in via generale nel precedente paragrafo 3, a cui è sufficiente rinviare, con riferimento a qualsiasi danno cagionato in generale a terzi da una AI utilizzata da uomini. Il tema, con riferimento specifico a danni cagionati dall’amministrazione telematica, è stato esaminato dalla più avveduta e lungimirante dottrina (75). Quest’ultima, non potendosi ipotizzare una soggettività dell’AI e prendendo atto delle difficoltà in punto di ascrivibilità sul piano psicologico in base all’art. 2043 c.c. (essendo invece evidenti e agevolmente provabili il danno e l’imputazione causale alla p.a. che adotti l’atto) del danno cagionato ad un singolo funzionario e alla stessa p.a. (ex art. 28 Cost.) a causa della complessità dell’elaborazione informatica di un provvedimento e della scarsa intellegibilità delle metodiche seguite da una AI da parte di chi la utilizza (oltre che della pluralità dei soggetti coinvolti in una gestione digitale della p.a.: produttore, programmatore, utilizzatore, ditte esterne utilizzate, manutentore, ecc.), auspica, in base ad una visione economica del diritto e a tutela del danneggiato (oltre che per escludere una irragionevole impunità della p.a.), che non può esser lasciato senza ristoro, l’utilizzo di forme di responsabilità oggettiva ex art. 2050 c.c., stante la “pericolosità” intrinseca dell’AI, bilanciabile (a tutela dell’utilizzatore p.a.), aggiungiamo noi, da polizze assicurative per responsabilità civile verso terzi.

In ordine, infine, alle conseguenziali ricadute in termini di *responsabilità amministrativo contabile* da uso di AI da parte delle pubbliche amministrazioni (o da privati legati alla p.a. da rapporto di servizio: es concessionari per la riscossione, laboratori e cliniche convenzionate, farmacisti ecc.), la tematica refluisce nei consueti binari del danno erariale diretto o indiretto devoluto alla Corte dei conti (76): il danneggiato da errori di una intelligenza

preimpostati (come fa invece l’algoritmo “tradizionale”) ma, al contrario, elabora costantemente nuovi criteri di inferenza tra dati e assume decisioni efficienti sulla base di tali elaborazioni, secondo un processo di apprendimento automatico (sulla base della distinzione tra la nozione di algoritmo e quella di intelligenza artificiale, il giudice amministrativo, ha, nel caso di specie, ritenuto che, al fine di ottenere la fornitura di un dispositivo medico con elevato grado di automazione non occorre che l’amministrazione facesse espresso riferimento a elementi di intelligenza artificiale, essendo del tutto sufficiente il riferimento allo specifico concetto di algoritmo, ben potendo quest’ultimo fornire un efficiente grado di automazione).

Su tale distinzione e relativi corollari, v. anche Tar Lombardia, Milano, Sez. II, 31 marzo 2021, n. 843.

(74) Tar Lazio, Roma, Sez. III-bis, n. 3769/2017, cit., che confuta la tesi contraria alla configurazione attizia del *software*. In terminis, Cons. Stato, Sez. VI, n. 2270/2019, cit., secondo cui “la regola tecnica che governa ciascun algoritmo resta pur sempre una regola amministrativa generale, costruita dall’uomo e non dalla macchina, per essere poi (solo) applicata da quest’ultima, anche se ciò avviene in via esclusiva”. La tesi in dottrina è condivisa da U. Fantigrossi, *Automazione e pubblica amministrazione*, Bologna, il Mulino, 1993, 51 ss., secondo cui “il software avrebbe natura di atto amministrativo generale per mezzo del quale l’amministrazione limita la propria discrezionalità, vincolando la propria azione rispetto ad una serie indeterminata di situazioni future”; A. Usai, *Le prospettive di automazione delle decisioni amministrative in un sistema di teleamministrazione*, in *Dir. informazione e informatica*, 1993, 163, che lo qualifica nei termini di atto amministrativo interno, mentre A. Masucci, *L’atto amministrativo informatico*, cit., 47 ss., inquadra il programma quale atto generale a contenuto non normativo attraverso il quale l’autorità amministrativa indirizza, predeterminandone modalità e contenuti, il proprio agire. Parla di atto amministrativo ad elaborazione elettronica F. Saitta, *Le patologie dell’atto amministrativo elettronico e il sindacato del giudice amministrativo*, in <www.cesda.it>, luglio 2003.

Contra, A.G. Orofino, *La patologia dell’atto amministrativo elettronico. Sindacato giurisdizionale e strumenti di tutela*, in *Foro amm.-CdS*, 2002, 2256, che ha escluso la qualificazione del programma informatico come atto amministrativo evidenziando come il *software* non sia materialmente imputabile al titolare dell’organo e che esso sia redatto in una lingua incomprensibile a chiunque non sia un tecnico, in violazione delle garanzie procedurali a favore dei cittadini interessati.

Interessante la tesi propugnata da A. Biamonte, *L’amministrazione digitale e la digitalizzazione procedimentale. Verso un nuovo esercizio della funzione. Problematiche e prospettive*, in F. Liguori (a cura di), *Studi sul procedimento e sul provvedimento amministrativo nelle riforme del 2005*, Bologna, Monduzzi, 2007, 252, il quale ha sostenuto che il *software* integreerebbe una mera modalità di esecuzione di una volontà amministrativa già espressa a monte della traduzione algoritmica. Sul tema, v. anche C. Napoli, *Algoritmi, intelligenza artificiale e formazione della volontà pubblica: la decisione amministrativa e quella giudiziaria*, in <www.rivistaaiic.it>, 28 luglio 2020.

(75) L. Viola, *L’intelligenza artificiale nel procedimento e nel processo amministrativo*, cit.; A. Masucci, *L’atto amministrativo informatico*, cit., 13; Id., voce *Atto amministrativo informatico*, in *Enc. dir.*, Aggiorn., vol. I; Id., *Procedimento amministrativo e nuove tecnologie*, cit., 93-94. Anche F. Morollo, op. cit., 27, rileva come “si potrebbe pensare ad un modello di responsabilità per i danni provocati nello svolgimento di un procedimento elettronico sganciato dal requisito della colpa e fondato unicamente su elementi oggettivi: in particolare, il criterio di riferimento potrebbe essere il rischio. Il rischio informatico”.

(76) Sulla responsabilità amministrativo-contabile, anche per danno indiretto, è sufficiente il rinvio a V. Tenore (a cura di), *La nuova Corte dei conti: responsabilità, pensioni, controlli*, Milano, Giuffrè, 2022, 4^a ed., 26 e 301. Sul tema N.M. Tritto, *Il danno erariale da intelligenza artificiale*, in <www.federalismi.it>, 20 marzo 2024; D.U. Galetta, J.G. Corvalán, *Intelligenza Artificiale per una pubblica amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in <www.federalismi.it>, 6 febbraio 2019.

artificiale utilizzata da una p.a. (ad esempio un errore diagnostico o interventistico in una struttura sanitaria pubblica acclarati in sede civile; una illegittima esclusione da un concorso o da una gara accertati dal g.a.; una erronea assegnazione di sede che porti ad un ricorso vittorioso per il lavoratore e danni risarciti dalla p.a.; un mancato rilascio di provvedimenti dovuti annullato dal g.a.; un recupero somme fondato su calcoli errati di un algoritmo acclarati dal g.a.; una erogazione di somme pubbliche sulla base di una istruttoria algoritmica fallace; un errore computazionale alla base di un minor introito tributario o concessorio, ecc.) porterà a danni diretti (da danno emergente o lucro cessante), o a danni indiretti nascenti da una previa azione civile (innanzi al g.o. o al g.a. a seconda dei casi) per risarcimenti a terzi lesi da enti pubblici. Dopo la condanna in giudicato dell'ente convenuto (frutto di un complesso giudizio civile per le ragioni istruttorie e probatorie sopra sunteggiate), ove non vi sia una polizza assicurativa a tutela della p.a. che abbia saldato il danneggiato integralmente, si configurerà un danno erariale, anche per le spese di lite sopportate dal soggetto pubblico soccombente, devoluto alla Corte dei conti, che dovrà vagliarlo alla stregua dei più ampi e concreti parametri della l. 14 gennaio 1994, n. 20 (come temporaneamente limitata dal c.d. scudo erariale dell'art. 21, d.l. n. 76/2020 e successive proroghe e come potrebbe essere modificata dal discusso “disegno di legge Foti” n. 1621/2023 presentato alla Camera e attualmente in accesa discussione), per ascrivere il danno alla colpa grave (senza spazi applicativi per forme di responsabilità oggettiva) del soggetto responsabile (costruttore, programmatore, utilizzatore, manutentore, custode, ecc.), purché pubblico dipendente o soggetto (anche persona giuridica) legato alla p.a. da rapporto di servizio.

Per completezza sul quadro contenzioso occasionato da AI, va da ultimo fatto cenno a qualche pronunciamento della Corte di giustizia Ue (77) sulle intelligenze artificiali utilizzate, sulla scorta della direttiva 2016/681 del Parlamento europeo e del Consiglio, del 27 aprile 2016, nell'uso dei dati del codice di prenotazione (PNR) a fini di prevenzione, accertamento, indagine e repressione dei reati di terrorismo e altri reati gravi. I dati del PNR (*Passenger name records*) sono riconosciuti tra le categorie più sensibili di informazioni personali. Si tratta dei dati compilati dalle agenzie di viaggi, dai vettori aerei e dai *tour operator*, e che contengono diverse informazioni relative ai passeggeri, tra le quali anche le condizioni mediche e le disabilità, oltre alle preferenze sui pasti, ovviamente i mezzi di pagamento, l'indirizzo di lavoro, l'indirizzo IP se si prenota *online* e le informazioni personali dei contatti di emergenza.

4. Sulla necessità di non abdicare, anche in giudizio, alla primazia dell'intelligenza non artificiale, ma naturale in quanto umana. Dubbi su una sentenza algoritmica

Senza addentrarci in ragionamenti filosofici, politici, sociologici o antropologici, che esulano dagli intenti di questo scritto (78), si vuole solo in conclusione sottolineare come l'uso di strumenti di lavoro che agevolino l'uomo nella vita, o pungolino la riflessione e favoriscano la risposta ad interrogativi, è un dato naturale e irrinunciabile: il fuoco, le armi da caccia, la ruota, gli strumenti primordiali per trasformare materie prime in alimenti ed

(77) Corte giust., Grande Sezione, 21 giugno 2022, n. 817, secondo cui (punti 194 e 195 della decisione), posto che nella direttiva Pnr, i criteri per trattare i dati Pnr devono essere “prestabiliti”, “194. [...] tale requisito osta all'utilizzazione di tecnologie di intelligenza artificiale nell'ambito di sistemi di autoapprendimento (machine learning), che possono modificare, senza intervento e controllo umani, il processo della valutazione e, in particolare, i criteri di valutazione sui quali si fonda il risultato della applicazione di detto processo nonché la ponderazione di tali criteri.

195. Occorre aggiungere che il ricorso a siffatte tecnologie rischierebbe di privare di efficacia l'esame individuale dei riscontri positivi nonché il controllo di liceità richiesto dalle disposizioni della direttiva Pnr. Infatti, come rilevato in sostanza dall'avvocato generale al paragrafo 228 delle sue conclusioni, tenuto conto dell'opacità che caratterizza il funzionamento delle tecnologie di intelligenza artificiale, può risultare impossibile comprendere la ragione per la quale un dato programma sia arrivato ad un riscontro positivo. In tali circostanze, l'uso di siffatte tecnologie potrebbe privare gli interessati anche del loro diritto a un ricorso giurisdizionale effettivo sancito dall'articolo 47 della Carta, che la direttiva Pnr mira, ai sensi del suo considerando 28, a garantire a un livello elevato, in particolare per contestare il carattere non discriminatorio dei risultati ottenuti”.

(78) È sufficiente rinviare, per considerazioni di più ampia portata sociale e politica, al recente saggio H. Kissinger, E. Schmidt, D. Huttenlocher, *op. cit.* Lo studio dell'ex segretario di Stato americano Henry Kissinger, l'ex amministratore delegato di Google Eric Schmidt e l'informatico e decano del Mit Daniel Huttenlocher evidenzia come presto l'umanità si ritroverà a imboccare un sentiero molto pericoloso, poiché l'IA sta cambiando il pensiero, la conoscenza, la percezione, la realtà e, di conseguenza, il corso della storia. Nonostante la velocità con cui avanza e progredisce, l'IA non è infatti governata da principi e concetti morali che la contengano e le diano dei limiti, sicché la sua rivoluzione può assumere pieghe inaspettate e condurre a esiti imprevedibili. Interrogandosi sui prossimi scenari possibili, tre fra i pensatori più autorevoli e lucidi di oggi riflettono così sull'intelligenza artificiale e su come stia trasformando il nostro modo di sperimentare la realtà, la politica e le società in cui viviamo.

Anche lo studio di A. D'Aloia (a cura di), *op. cit.*, evidenzia come l'intelligenza artificiale, anche in combinazione con le tecnologie informatiche e la genetica, sembra davvero delineare i contorni di un “mondo nuovo” piuttosto che essere semplicemente un insieme di risorse tecnologiche aggiuntive e avanzate rispetto a quelle già consolidate. Tecnologie “profonde”, le definisce Jason Preston, “invenzioni che ci reinventano”, che “segnano un genere diverso di cambiamento del mondo” proprio perché puntano a “progettare meglio l'evoluzione”, e a “reinventare la natura”. Un “nuovo mondo” che contiene e propone interrogativi profondi e inediti, i quali investono i campi e le categorie più qualificanti della civiltà sociale (e giuridica: salute, istruzione, lavoro, procedure democratiche, sicurezza, giustizia), e arrivano a prefigurare scenari (con tratti persino inquietanti) in cui può venire in discussione l'essenza stessa di ciò che consideriamo identità umana.

altri oggetti, il motore a scoppio, i mezzi di trasporto (di terra, mare e cielo), la scrittura e la stampa, le catene di montaggio, le macchine calcolatrici, sono stati nei secoli sviluppati o sostituiti da strumenti di lavoro e di ricerca sempre più raffinati, diversificati e complessi, per venire incontro alle crescenti esigenze umane e per dispensare l'uomo dal "peso della fatica" e, almeno sulla carta, per migliorare la sua "felicità".

Tali strumenti hanno infatti aiutato l'uomo a migliorare la qualità della vita, ad allungarla, a ridurre tempi di lavorazione, ad incrementare la produttività, a limitare errori umani, ad attenuare il logorio fisico in lavori usuranti o alienanti, a favorire una crescita culturale omogenea sul territorio, a fornire in modo capillare servizi un tempo appannaggio di pochi. Le enormi e diversificate attività produttive, gestionali, ludico-ricreative e persino sociali espletabili anche restando a casa, o in una amena località montana o marina utilizzando un PC o uno *smartphone*, sono sotto gli occhi di tutti.

Ma sino ad oggi il ruolo degli strumenti di lavoro e dei mezzi volti a fornire beni e servizi ai cittadini è rimasto ancillare e servente rispetto al prioritario uso del dono più prezioso della vita, ovvero *l'intelligenza umana* in tutte le sue manifestazioni: ideative, realizzative, affettive, emotive, programmatiche, elaborative, linguistiche, gestuali, ma anche etiche, di libero arbitrio e di buon senso nelle scelte quotidiane, personali e lavorative.

Tutte queste manifestazioni vitali dell'intelletto umano, frutto di empatia (con genitori, insegnanti, amici, compagni o coniugi, figli, clienti, pazienti, avvocati, colleghi), doti genetiche, fatica nello studio, esperienze di vita sono governate da un cervello organico che ha, in misura sicuramente mutevole (sotto il profilo quantitativo e qualitativo) tra gli individui, dei parametri e dei limiti fondamentali e innati: il buon senso, la coscienza, il senso del limite, il rispetto del prossimo e di se stessi, l'etica, la fiducia, l'amore, l'empatia, concetti fluidi e cangianti, ma sicuramente presenti, ancorché in graduazioni diverse, negli irripetibili e irripetibili esseri umani e non in un robot-antropomorfo e nei suoi più profondi algoritmi.

Quest'ultima è la tendenziale connotazione dell'essere umano rispetto alle *res*, alle cose (anche se alcune *res*, quali gli animali e in misura minore le piante, presentano innegabilmente alcune connotazioni umane, ed alcuni esseri umani presentano caratteristiche primordiali che li avvicinano alle *res*) (79). Dunque, non ha alcun senso tentare di antropomorfizzare i modelli di conversazione o i testi partoriti da LaMDA o Chat-GPT, che non sono soggetti senzienti.

E allora, se il crescente utilizzo di macchine intelligenti, innegabilmente utili e necessarie in molti settori della vita (sanità, servizi, trasporti, logistica, ricerca, assistenza), potrà portare, nell'immediato futuro, all'accantonamento progressivo di molte basilari manifestazioni creative ed elaborative dell'intelligenza umana, sostituita da una sempre più sofisticata intelligenza artificiale, ciò significherà abdicare progressivamente alla funzione più connotante dell'essere umano: *il pensare*, in tutte le sue vastissime manifestazioni e con tutte le sue pregresse esperienze non solo tecniche, giuridiche, ma anche e soprattutto umane e sentimentali. E ciò porterà, forse, ad una crisi sociale ed etica e a dubbi esistenziali sul "senso della vita" ove delegata a "macchine". L'AI potrebbe infatti dispensare l'Uomo non più dalla fatica e dalla *routine* di alcuni lavori, ma, ancor più in radice, dal partorire pensieri, la cosa più immateriale ma preziosa che esista!

Che spazio residuerà allora, a fronte di macchine sempre più intelligenti e tra loro connesse (sino a raggiungere la "singolarità" ovvero il presunto momento in cui le macchine diverranno più intelligenti dell'uomo) (80), per la dignità e la responsabilità umana, per il libero arbitrio, per la coscienza, per l'intelligenza dell'uomo? Probabilmente si arriverà alla creazione di quello che il neuroscienziato Michel Desmurget ha definito il "*cretino digitale*" (81), "evoluto" da *homo sapiens* ad *homo numericus*.

A tale prioritario ed esistenziale problema se ne aggiungono altri: a) qualsiasi tecnologia, e dunque una AI, con i suoi algoritmi, può essere utilizzata a fin di bene ma anche a fini malvagi, utilizzando nostri dati personali e

(79) Sulla differenza tra intelligenza artificiale e intelligenza umana si sofferma anche E.J Larson, *Il mito dell'Intelligenza artificiale*, Milano, Angeli, 2022, secondo cui l'intelligenza artificiale lavora sul ragionamento induttivo, analizzando *dataset*, enormi quantità di informazioni, per prevedere possibili risultati di una scelta o di un'azione. Gli esseri umani, invece, formulano congetture che tengono conto del contesto e dell'esperienza. L'intelligenza umana è una rete di ipotesi formulate sulla base di ciò che si conosce del mondo. «Non abbiamo la minima idea di come programmare quella forma di ragionamento intuitivo noto come "abduzione", che è a fondamento del senso comune», rileva Larson. E osserva: "questo è il motivo per cui Alexa non può capire ciò che le chiediamo e per cui l'intelligenza artificiale non può andare oltre certi limiti".

Nonostante le entusiaste tesi propugnate da cultori della umanizzazione e soggettivizzazione delle AI cui si è fatto cenno in nt. 47, le grandi reti neurali odierne producono risultati accattivanti che si avvicinano al linguaggio e alla creatività umani grazie ai progressi nell'architettura, nella tecnica e nel volume dei dati immessi nell'AI. Ma i modelli si basano sul riconoscimento del modello fornito dall'uomo, non sull'arguzia, sul candore, sull'intuito, sulla coscienza.

(80) In tale momento l'AGI (*Artificial general intelligence*), diverrà ASI (*Artificial superhuman intelligence*), capace di ricopiare se stessa, di aggiornarsi, migliorarsi e risolvere meglio dell'uomo taluni problemi (è quello che in realtà è già accaduto nella diagnostica e in parte della interventistica in medicina), creando una sudditanza dell'uomo nei confronti della macchina, come sta lentamente accadendo per le automobili ipertecnologiche, sulle quali è difficile intervenire persino per un rabbocco d'olio o di liquidi per tergicristalli.

(81) M. Desmurget, *Il cretino digitale*, Milano, Rizzoli, 2020, che evidenzia le ricadute dell'evoluzione tecnologica soprattutto sulle generazioni più giovani.

sensibili (come sta già accadendo tramite i *social media* che profilano ogni utente) (82) o hackerando e alterando algoritmi (83); b) la mancanza di buon senso in una AI potrebbe portare, leggendo freddamente i dati e gli *alert*, ad incidenti nucleari o guerre, non comprendendo la distinzione tra una esercitazione ed un attacco bellico, come invece accaduto negli anni '60 in epoca di guerra fredda Usa-Urss, in cui si evitò un conflitto grazie ad operatori umani dotati di buon senso; c) la crescita delle AI, dopo aver portato alla scomparsa di molti lavori manuali pesanti, noiosi o ripetitivi, sopprimerà anche preziosi lavori di assistenza alla persona (84) e ridurrà drasticamente ceti impiegatizi e liberi professionisti, creando disoccupazione; d) la robotica militare (droni) potrebbe avere portata devastante in punto di obiettivi da colpire senza priorità etica (sterminando popolazioni in aree ad alta urbanizzazione); e) le intelligenze artificiali creano talvolta correlazioni spurie e i dati acquisiti si possono tradurre in pregiudizi (85).

Dunque, a fronte di visioni spesso aprioristicamente contrapposte sul ruolo dell'AI nella attuale e futura società, come ben sintetizzato da Cosimo Accoto (86), di volta in volta e a seconda delle fazioni, la tecnologia digitale manifesta la sua potenza salvifica o, viceversa, la sua forza distruttiva. È accaduto per l'industrializzazione, sta accadendo per la digitalizzazione: *«la tecnologia sarebbe allora, insieme, “ruina” e “remedio” per riprendere due concetti che Machiavelli assegnava all'arte della politica: essere rovina e rimedio al tempo stesso. Pharmakon, dicono i filosofi, medicina e veleno insieme»*.

(82) Accedendo da un PC o uno *smartphone* a qualsiasi sito commerciale, e non solo, gli algoritmi di predizione diventano algoritmi di profilazione: non solo predicono quale sarà il nostro comportamento, ma ci inviano pubblicità su temi sui quali abbiamo fatto ricerche e producono un nostro comportamento, come sanno bene tutti gli esperti di *marketing*, per cui i dati profilati portano ad acquistare quanto suggerito o a prendere più di un prodotto a cui si era interessati.

(83) La *cybersicurezza* è un problema serio che in Italia il Governo sta affrontando creando una apposita Agenzia non casualmente vicina e contigua ai servizi segreti. Da parte poi di alcuni studiosi si propugna l'idea di una “AI amichevole” in quanto utile per l'umanità e sicura: i suoi algoritmi prevedibili, sicuri, non facilmente hackerabili e attendibili, in caso di fallimento, dovrebbero produrre danni delicati.

(84) Si pensi alle già segnalate badanti artificiali (v. il dispositivo citato in nt. 55) con cui anziani posso dialogare per ore raccontando all'infinito le stesse cose senza timori di irritare o infastidire l'interlocutore (in realtà tale ruolo è già assolto oggi da alcuni conduttori di trasmissioni di intrattenimento o a premi, banali e basiche, in TV, in fascia mattutina, pomeridiana e persino serale, divenuti più presenti di taluni parenti per tantissimi anziani soli, che li considerano componenti della famiglia o, addirittura, gli unici compagni in giornate di solitudine), o alle bambinaie-robot (incapaci però di far evolvere linguisticamente e sul piano esperienziale e sociale i piccoli, come invece accade in un cortile o in un parco giochi frequentati da esseri umani), ai computer psicoterapeuti capaci di identificare la depressione, ai docenti-robot di alto profilo che dequalificano insegnanti umani.

Sulla empatia tra uomo e intelligenza artificiale, v. l'interessante intervista al neuroscienziato Vittorio Gallese *L'intelligenza artificiale cambierà il nostro cervello, ma evitiamo scenari apocalittici*, in <www.huffingtonpost.it>, 19 luglio 2023, il quale, nel ricordare la profetica rappresentazione di tale empatia nel film *Her*, aggiunge “*ho visto persone parlare con una lavatrice, anche senza scomodare i sistemi di intelligenza artificiale. Un'altra caratteristica umana è proiettare la propria intenzionalità con tutto il portato di sentimenti ed emozioni sugli oggetti inanimati; quindi, è verosimile che ciò possa accadere*”. Lo studioso aggiunge che “*molti di questi algoritmi hanno già un modo di proporsi che si misura e cerca di adattarsi allo stile di relazione dell'umano*”.

Sulla scomparsa di alcuni lavori, sostituiti da AI, si sofferma J. Kaplan, *Le persone non servono. Lavoro e ricchezza nell'epoca dell'intelligenza artificiale*, Roma, Luiss University Press, 2021.

(85) In modo plasticamente evidente un attento studioso, A. Del Noce, *Giustizia predittiva. Il mondo forense si interroga su trasparenza, conformità e nuovi scenari di responsabilità professionale*, in <ntplusdiritto.ilsole24ore.com>, 15 febbraio 2023, formula diversi esempi sulle correlazioni spurie che si traducono in pregiudizio, evidenziando: “*Se lo dessi in pasto al migliore degli algoritmi di machine learning a mia disposizione e gli chiedessi chi è il prossimo presidente degli Stati Uniti l'algoritmo mi dice sicuramente un uomo, un po' meno sicuramente di etnia caucasica, un po' meno sicuramente tra 55 e 70 anni, perché non mi dice una donna perché i dati che non sono altro che le scelte passate non includono al loro interno nessuna donna. [...] Nel 2015 il software di Google ha categorizzato le persone di colore come... gorilla. Con Face up, applicazione di modifica dei volti, il filtro che doveva far apparire le persone più belle le faceva apparire più bianche. Fino a poco tempo fa, se su LinkedIn cercavate il nome di una donna, ad es. Francesca Rossi, il motore di ricerca suggeriva Francesco Rossi. Come dire che gli uomini andavano meglio delle donne. Anche un algoritmo utilizzato dall'Università della Virginia ha subito un'accusa di sessismo: quando caricavano scene di persone in cucina e si mostravano uomini in cucina, l'algoritmo diceva che erano donne. In molti stati degli Usa i giudici sono autorizzati ad usare un algoritmo denominato risk assesment che in base a circa 150 parametri che stabilisce la possibilità dell'imputato di commettere nuovi reati. I giornalisti di ProPubblica, una testata investigativa indipendente, hanno testato 7.000 di questi punteggi ed hanno verificato che le persone di colore vengono sistematicamente svantaggiate. Cioè l'algoritmo associava la pelle scura alla possibilità di delinquere. Non correlando i fattori socio-economici che determinano normalmente la delinquenza. In altre parole, il problema è che i dati registrano quelle che sono le scelte: se le scelte sono ingiuste i dati possono essere pure perfetti ma al loro interno per criteri del passato e criterio ingiusti perpetrano quelli che sono dei criteri logici che chiamiamo pregiudizi. Cioè i dati diventano pregiudizi*”.

(86) C. Accoto, *Il mondo in sintesi*, Milano, Egea, 2022. Del filosofo digitale v. anche Id., *Il mondo ex machina*, Milano, Egea, 2019. In ambo i libri lo studioso si chiede se il cambiamento evocato dall'era della simulazione digitale sia trasformativo a beneficio dell'umanità o sia piuttosto distruttivo della vita su questo pianeta e fornisce articolate risposte senza estremismi aprioristici.

I robot senzienti hanno ispirato decenni di fantascienza distopica. Oggi, come si è detto, la vita reale ha iniziato ad assumere una sfumatura fantastica con GPT-3, un generatore di testo in grado di partorire un parere legale come la sceneggiatura di un film, e DALL-E 2, un generatore di immagini in grado di evocare elementi visivi basati su qualsiasi combinazione di parole: entrambi dal laboratorio di ricerca OpenAI.

Va inoltre rimarcato, sotto altro profilo, come il legame sempre più intenso tra uomo e animale e tra uomo e intelligenze artificiali è sicuramente sintomatico di una evoluzione (o una involuzione) della società e dei rapporti interpersonali tra esseri umani, che sempre di più preferiscono la semplicità di un animale o l'accondiscendenza ragionata di una AI (a cui delegare sempre più compiti manuali e cerebrali propri dell'uomo, quali anche il crescere ed accudire figli o prendersi cura di genitori anziani) alla complessità delle relazioni interpersonali o dell'agire, dello studiare, del lavorare e del pensare, tipici dell'essere umano. E la preferenza verso un animale o una AI (ed il suo mondo virtuale fatto anche di giochi, intrattenimenti, dialoghi non umani), mere *res* seppur intelligenti, rispetto ad un uomo o una donna reali, deve far riflettere anche sulla solitudine, l'egoismo, il disimpegno che connotano sempre più di frequente la società.

Elaborare un quadro, una canzone (87), una poesia, un racconto, una fiaba, una sentenza, una memoria, un contratto, un provvedimento e persino un amico (88) o un'amante virtuale, come la filmografia più lungimirante ha da tempo profetizzato e raccontato (89), attraverso una macchina sofisticata (che comunque attinge a pregresse opere dell'ingegno umano rielaborandole), significa rinunciare all'essere vivi e vitali, preferendo la stasi contemplativa dell'opera, seppur stimolante, di una macchina, ovvero una *res intelligente*, priva di sentimenti, coscienza, etica e buon senso, di amore e gratitudine, che connotano qualsiasi opera e manifestazione dell'essere umano. Senza dimenticare che un robot intelligente, pur potendo comporre musica, dipingere, scolpire, redigere un saggio, rendere un parere legale, *genera parole senza pensare* e non sa se il suo prodotto sia bello e cosa trasmetta sul piano emotivo, perché privo di sensibilità. Le parole e le immagini generate dai sistemi di intelligenza artificiale come LaMDA o Chat-GPT producono risposte basate su ciò che gli umani hanno già pubblicato su Wikipedia,

(87) V. la successiva nt. 101.

(88) Diverse sono le "applicazioni emotive", tese a far provare intimità, connessione umana, affetto, amicizia, reperibili online.

Il programma *Replika*, nato grazie ad Eugenia Kuyra, che aveva perso nel 2015 in un incidente il suo miglior amico e voleva replicarlo virtualmente in forma di *chatrobot*. Oggi *Replika* vanta 10 milioni di utenti che cercano un compagno virtuale, elaborare traumi, ansia, violenze testare il proprio orientamento sessuale, intraprendere giochi erotici. Il programma risulta bloccato in Italia dal Garante della *privacy* in quanto pericoloso per i minori.

Decisamente più votato a fantasie erotiche è *Blush*, una app molto audace che si avvicina ad un simulatore di Tinder. Ma offrono aiuti in vari campi "emotivi" diversi altri programmi per persone emotive, timide o sole: *Rizz*, *Your-Move AI*, *Personal AI*, *Keys AI*, *Elate*, *Teaser AI*.

Tali AI, capaci (in modo ancora imperfetto) di riconoscere le emozioni dell'utente (monitorando il respiro, la risposta galvanica della pelle, l'intonazione e velocità della voce, espressioni facciali, ecc.) e di reagire in modo apparentemente emotivo, pongono problemi etici molto seri nell'intraprendere "amicizie" con persone emotivamente bisognose.

Come è noto, le emozioni sono e restano umane in quanto prodotte da neuromodulatori diffusi solo nel cervello umano, circostanza che esclude una AI con affetti, anche se studi e sperimentazioni recenti si sono incentrati su simulazioni di neuromodulazione nel produrre "computer da compagnia" per anziani, disabili, bambini e neonati o per amanti di "giochi per adulti" interattivi: computer badanti, bambinaie o partner sessuali.

(89) Indimenticabile, e quasi profetico, è il film fanta-sociologico di e con Alberto Sordi *Io e Caterina*, che, nel lontano 1980, racconta di Caterina, un robot tuttora dalle fattezze femminili che sbriga le faccende domestiche meglio di una persona in quanto non dorme e non ha bisogno di nutrirsi (aggiungiamo noi che non reclama stipendi, contributi né fa cause di lavoro) e che palesa progressivamente l'attaccamento, l'apprensione, la gelosia (prelude al protagonista Enrico-Alberto Sordi relazioni con altre donne) tipiche di una compagna umana, sfuggendo alla iniziale programmazione dell'ingegnere che aveva venduto la "macchina" ad Enrico. Caterina detta le sue condizioni a Enrico: lei lo servirà devotamente e senza protestare solo se lui le offrirà il rispetto e la dedizione che chiederebbe una donna vera.

In tempi più recenti e con maggior legame con l'attuale società iperteleumatica, il film *Lei (Her)* del 2013, scritto e diretto da Spike Jonze, con protagonista Joaquin Phoenix (film, che si è aggiudicato il premio Oscar per la miglior sceneggiatura originale) descrive un futuro prossimo, non troppo lontano, nel quale i computer hanno un ruolo di primissimo piano nella vita anche affettiva delle persone. Difatti, con l'uscita sul mercato del nuovo sistema operativo OS 1, provvisto di intelligenza artificiale in grado perfino di apprendere ed elaborare emozioni (quali la gelosia verso una donna reale o simulare rapporti fisici), si crea un legame empatico, poi affettivo, infine sentimentale tra il protagonista Theodor (uomo in procinto di separarsi) e Samantha, donna telematica creata dal programma OS 1. Tuttavia, Samantha confessa dopo del tempo a Theodor che l'enorme velocità di elaborazione e di evoluzione delle intelligenze artificiali aveva portando lei e i suoi simili (che avevano intrapreso a loro volta relazioni con esseri umani) sempre più lontano dalla percezione umana e le è sempre più difficile riconoscersi nel rapporto con l'uomo-Theodor. I due protagonisti tristemente si dicono dunque addio e lei scompare definitivamente dal computer di Theodor che torna dalla sua ex moglie.

Degna di menzione è anche, sul piano cinematografico, la nota scena di sesso in *Blade Runner 2049*, che coinvolge il personaggio di Ryan Gosling, il suo ologramma Joi e una prostituta di nome Mariette, ingaggiata per colmare il divario fisico tra i due amanti.

Da ultimo un distruttivo utilizzo di una invasiva AI (capace di penetrare e dominare sistemi telematici di governi, servizi segreti, forze di polizia, computer comuni) da disinnescare con una ambita chiave meccanica, si incentra l'ultima tappa della saga dei film *Mission Impossible: dead reckoning*, del 2023, con il maturo ma sempreverde Tom Cruise.

Reddit, bacheche e ogni altro angolo di Internet, ma questo non significa che il modello ne comprenda il significato.

E ciò vale anche per una *sentenza* redatta da una macchina (90) (o per una *conciliazione* giudiziale espressiva di Adr) (91) che, pur dotata di poderosa memoria storica fatta di dati normativi, di milioni di precedenti giurisprudenziali e di scritti dottrinali, resta priva di esperienze umane (*in primis* dell’“esperienza del dubitare”), di relazioni interpersonali, di gioie e di dolori, di sentimenti, emozioni e, soprattutto, difetta di coscienza, etica e di buon senso, concetti impalpabili che si apprendono (non per tutti) *vivendo nella vita reale e non in quella degli algoritmi*, esperienza che una intelligenza artificiale, purtroppo per lei, non potrà mai provare, anche se si coinvolgesse (come in parte sta già accadendo) umanisti e teologi nei gruppi di lavoro che elaborano macchine intelligenti ed algoritmi, nel tentativo di fissare alcune regole di “*algoretica*”, ovvero algoritmi etici (92) volti a creare robot antropomorfi.

Sul piano elaborativo di un testo, e dunque di una sentenza, i segnali elettrici di una intelligenza artificiale non possono che produrre altri segnali elettrici o conseguenze fisiche come forza e movimento, ma non sensazioni, dubbi e sentimenti che sono qualitativamente diversi: è la *coscienza* (non casualmente valorizzata in tutti i codici deontologici delle libere professioni che invitano ad agire “*secondo scienza e coscienza*”: v. art. 5 del codice deontologico dei medici), che capisce la situazione concreta e la fa capire ad un Uomo, e ad un Uomo-giudice: e questa l’essenza della differenza tra un incosciente robot e un Uomo. Una semplice intelligenza artificiale (93) non può dunque sostituire il “guizzo”, la *sagitta*, la scintilla umana fatta di intuito, spirito, amore, sensibilità, coscienza e comprensione del fatto concreto, soprattutto in un processo.

(90) Sul tema, L. Nazzicone, *La motivazione della sentenza. Stile e tecniche*, Milano, Giuffrè, 2024, 210 ss., che si sofferma sulla decisione robotica; Aa.Vv. (a cura di A. Nuzzo e G. Olivieri), *Algoritmi se li conosci, li regoli*, in *Analisi giur. economia*, 2019; M. Luciani, *La decisione giudiziaria robotica*, E. Vincenti, *Il problema del giudice-robot*, e M.R. Covelli, *Dall’informatizzazione della giustizia alla “decisione robotica”*, il giudice del merito, tutti e tre in A. Carleo (a cura di), *Decisione robotica*, Bologna, il Mulino, 2019, rispettivamente 85 ss., 122 ss. e 125 ss.; F.C. Gastaldo, *Il giudice-robot: l’intelligenza artificiale nei sistemi giudiziari tra aspettative ed equivoci*, in <www.iusinitinere.it>, 22 marzo 2021; M. De Felice, *La macchina della decisione*, Torino, Aragno, 2021.

V. anche E. Calzolaio (a cura di), *op. cit.*, secondo cui uno degli aspetti più rilevanti della diffusione degli impieghi dell’intelligenza artificiale è che con essa sembra incrinarsi, se non proprio spezzarsi, uno dei cardini del pensiero giuridico moderno, costituito dalla possibilità di imputare una condotta ad un soggetto in base alla autonomia del processo decisionale che ne è alla base. Sul tema e sulle ricadute dell’uso di AI per le professioni legali (magistratura, notariato, avvocatura, docenza universitaria), v. L. Cappello *et al.*, *op. cit.*

(91) Sull’applicabilità dell’AI alle metodiche Adr si rinvia alla precedente nt. 11. Sul tema, C.E. Bruno, *op. cit.*

(92) Osserva l’Accademia della Crusca nel suo sito che l’*algoretica* rappresenta lo studio dei problemi e dei risvolti etici connessi all’applicazione degli algoritmi. Trattasi di parola macedonia formata da *algor*, abbreviazione di *algoritmo* (“insieme di regole per la risoluzione di un calcolo numerico” e per estensione “metodo o procedimento matematico per la risoluzione di un problema”) ed *etica* (“complesso delle norme morali e di comportamento pubblico e privato proprie di un individuo o di un gruppo”).

Tra le sue prime attestazioni vi è lo studio di P. Benanti, *Oracoli. Tra algoretica e algocrazia*, Bologna, Sossella, 2018, secondo il quale “*Le implicazioni sociali ed etiche delle AI e degli algoritmi rendono necessaria tanto un algor-etica quanto una governance di queste invisibili strutture che regolano sempre più il nostro mondo per evitare forme disumane di quella che potremmo definire una algo-crazia*”.

Anche L. Floridi, *Etica dell’intelligenza artificiale*, Milano, Cortina, 2022, individua i cinque principi fondamentali che determinano il carattere etico dei sistemi di intelligenza artificiale, e sono: *beneficenza*, *non malevolenza*, *autonomia*, *giustizia* ed *esplicabilità*. E aggiunge che è proprio la mancanza di questo ultimo principio che continua a compromettere l’eticità di certe tecnologie, quando non condividono informazioni sul proprio algoritmo, sul *dataset* di riferimento o sui metodi di addestramento. Se non c’è un’adeguata trasparenza non è possibile indagare sugli scopi e sui pericoli che si possono celare, oggi o domani, dietro questi sistemi.

Sull’etica delle tecnologie va poi segnalato l’interessante studio del teologo francescano P. Benanti, *Digital age*, Cinisello Balsamo, San Paolo, 2020, che approfondisce i profondi e veloci cambiamenti in corso, le caratteristiche della cultura contemporanea, le nuove coordinate esistenziali ed affronta alcune questioni molto rilevanti quali i nuovi valori, il rapporto tra la tecnologia e la qualità della vita di questa nuova epoca digitale, comprese le sfide del “postumano”, le intelligenze artificiali e l’algor-etica.

(93) Un programma che muove una macchina è fatto da una sequenza di comandi ideati da una mente umana e scritti in linguaggio booleano che usa due sole lettere: 0 ed 1, espressive di logica binaria costituita da *bit* (vero-falso, alto-basso, esiste-non esiste, questo-quello). Invece il comportamento quantistico di singoli atomi e molecole all’interno di una cellula vivente dà origine ad un ordine dinamico molto più complesso e non ancora ben compreso, che non può essere descritto dalle leggi quantistiche statiche che regolano i comportamenti di oggetti, quale è una AI. Smontando una macchina è possibile rimontarla e ricostruirla identica: se si potesse smontare una cellula in movimento e viva, sarebbe impossibile rimontarla o rimontarla uguale a prima.

Una macchina può essere accesa e spenta dall’uomo tantissime volte (si pensi ad un PC ed alla sua memoria), mentre un organismo vivente si spegne una sola volta: quando muore.

A differenza di una macchina, un organismo vivente non rimane mai la stessa entità fisica e psicologica da un istante all’altro.

Sulle 7 differenze tra una cellula e un robot, si rinvia a F. Faggin, *op. cit.*, 126.

A ciò aggiungasi che la tecnica di “*giustizia predittiva induttiva*” seguita dagli algoritmi, e fondata sulla valutazione dei precedenti prevalenti (e dunque scarsamente utilizzabile per casi nuovi o poco esplorati) (94), porta ad una evoluzione, anche nel Paese di *civil law*, verso la metodica del “precedente vincolante” tipico dei Paesi di *common law*, snaturando le peculiarità del nostro ordinamento giudiziario, in cui la forza di un precedente è solo nella sua autorevolezza e capacità di convincimento del singolo giudice, fondata sulla condivisibilità o meno del ragionamento seguito. Ma, soprattutto, vi è il grave rischio della standardizzazione decisoria: difatti, se si ritiene che una causa abbia un basso livello di successo perché contraria a molti precedenti, allora nessuno proporrà tale causa, con la conseguenza di frustrare la spinta naturalistica all’evoluzione del diritto (ad esempio, la sentenza n. 500/1999 delle Sezioni unite sul risarcimento dell’interesse legittimo non sarebbe mai stata partorita da una AI, ancorata a meri precedenti, tutti ostativi). Nella decisione finale di un giudice, come ben sa chi abbia affrontato camere di consiglio e stesure di sentenze, non assume rilevanza il numero di precedenti a favore o contro una soluzione, ma unicamente la correttezza degli argomenti esposti *pro* e *contro*.

Inoltre, se l’AI viene utilizzata come “scorciatoia” per evitare di lavorare cerebralmente alla stesura di una sentenza, di una comparsa difensiva (95), di una conciliazione, di un rogito notarile, di un provvedimento amministrativo, proseguirà, anche sul lavoro, il crescente abbassamento del livello culturale di categorie “colte” e “pensanti”, quali magistrati, avvocati, notai, dirigenti pubblici, in perfetta sintonia con il generalizzato abbassamento culturale e linguistico dei giovani nascente dall’uso sistemico di cellulari e PC per redigere temi e ricerche a scuola, risolvere problemi matematici, preparare esami universitari (tra l’altro in università telematiche e, in qualche caso “telepatiche” per la pochezza selettiva tra gli studenti negli esami), rinunciando al peso della “*fatica dello studio*”.

E dunque, solo chi ha sperimentato e vissuta la *fatica dello studio* (liceale, universitario e postuniversitario) prima e *della fatica del decidere* poi, solo chi, oltre ad aver letto gli atti di causa e ad aver consultato precedenti giurisprudenziali (come può certamente fare, ed anche meglio, una attenta e veloce AI), solo chi si è confrontato, alla luce del proprio bagaglio umano e professionale, con avvocati e colleghi in migliaia di discussioni in aula e in camera di consiglio, solo chi ha patito la “sofferenza del capire e del dubitare”, solo chi ha letto il volto, le voci, le posture e le espressioni di imputati, di parti in causa, di avvocati e di pubblici ministeri, solo chi ha odorato e percepito gli umori in un’aula di udienza, un corridoio di un tribunale, una caserma dei Carabinieri e un’aula penitenziaria, solo chi è capace di confrontarsi con la propria coscienza prima di assumere una qualsiasi decisione, è in grado di redigere una sentenza degna di questo nome, ovvero il frutto di un’operazione valutativa molto, molto più complessa di quella di cui è capace una “banale” intelligenza artificiale.

Il complesso giudizio tra esseri umani, a nostro avviso, non può che seguire parametri e criteri umani e non quelli, esclusivamente logici, di una macchina, ancorché intelligente in quanto creata da uomini e ancorata a criteri (logici) e conoscenze immesse nella macchina dall’uomo e dalla sua mente creativa.

Non bisogna dunque rinunciare alla funzione essenziale dell’Uomo-magistrato: studiare, inquadrare, valutare e riflettere con equilibrio, competenza e indipendenza, sulle tante patologie comportamentali che la realtà demanda ad un tribunale. Non si può dunque abdicare alla sapienza e alla complessità, *esclusivamente umana*, del decidere: una macchina, per quanto sofisticata essa sia, *non vive* e non ha *coscienza*; quindi, non può giudicare

(94) L’impostazione basata su meri calcoli statistici dei precedenti giurisprudenziali ha una portata limitata ai soli casi in cui ci siano numerosi precedenti, così da escludersi i casi più complessi involgenti novità normative, non ancora oggetto di stratificati orientamenti giurisprudenziali.

Inoltre, la giustizia predittiva fondata su precedenti presenta un altro rischio: quello della fallacia del precedente, in quanto un errore giurisprudenziale ripetuto tante volte, non può divenire, per ciò solo, un “non errore”. Se una sentenza è errata, allora vi è il rischio che venga seguita solo perché precedente giurisprudenziale. E ciò è ancor più pericoloso se il precedente è delle Sezioni unite o delle Sezioni riunite.

(95) Cambierà la stessa pratica per gli avvocati: se un tempo si basava sul ragionamento giuridico, sulla ricerca manuale condotta con le banche dati, sullo studio della pratica, sulla raccolta delle informazioni, in futuro questa parte potrebbe essere affidata all’AI, mentre l’intervento dell’intelligenza umana sarà dedicato alla scelta di cosa chiedere all’AI, di come interpretare i risultati forniti e quale strategia intraprendere sulla base delle previsioni fatte dall’AI.

uomini e relative azioni e interazioni (96), né avere contezza delle contestualizzazioni dei fatti e delle occasionali c.d. illegalità praticabili di Niklas Luhmann (97).

Forse, ma si ribadisce il *forse*, solo questioni minimali, poco complesse, connotate da modestissimi margini valutativi e discrezionali, o questioni seriali, possono essere delegate ad “oggetti pensanti”, ovvero a “giudici-robot” espressivi di AI (98), per accelerare i tempi della giustizia, affogata da tanti contenziosi temerari, vili o pretestuosi, che non meritano neppure il vaglio di un giudice umano, lo spreco di energie cerebrali umane, e che una buona macchina potrebbe decidere in tempi rapidi, magari condannando il soccombente anche alle spese... per lite temeraria!

Ovviamente quest’ultima opzione, favorevole ad una sentenza “non complessa” redatta esclusivamente da una intelligenza artificiale, pone problemi di compatibilità di tale approdo con il dettato costituzionale (art. 102 Cost.), che devolve alla magistratura “in carne, ossa e cervello” la funzione giurisdizionale e pone anche un problema ontologico ancor prima che qualificatorio: se il frutto cerebrale di una AI-giudicante sia giuridicamente definibile *sentenza in nome del Popolo Italiano*, conclusione che riteniamo di escludere, sebbene la giurisprudenza, come sopra ricordato (99), abbia ritenuto utilizzabili algoritmi per la redazione di atti amministrativi meno complessi ed abbia qualificato come “atto amministrativo” il *software*. La sentenza, per esser tale, non può invece che essere frutto di una attività cerebrale umana assolutamente ben prevalente.

(96) La nostra tesi, per qualsiasi AI operante anche in altri campi, è condivisa da molti. Ad esempio, M. Tampieri, *op. cit.*, rimarca la necessità di un governo antropocentrico dell’intelligenza artificiale, da declinare in chiave personalistica e solidaristica, al fine di tutelare le libertà e i diritti fondamentali di tutti coloro che interagiscono con i sistemi di IA. Dunque, secondo l’A., si rivela importante un approccio *human in command* che permette di mantenere il controllo dell’uomo sulle decisioni della macchina intelligente, il cui margine di rischiosità è sì particolarmente ridotto nelle macchine dotate di un elevato livello di autonomia, ma non è comunque completamente eliminabile.

Anche R. Cucchiara, *L’intelligenza non è artificiale*, Milano, Mondadori, 2021, rimarca come le nuove tecnologie sono sì delle macchine di calcolo potentissime e velocissime, ma la loro cosiddetta “intelligenza” è anche “frutto del pensiero, del controllo e del comportamento umano. Siamo noi i suoi maestri e i suoi giudici. Quindi saranno il nostro sentire, la nostra cultura e il nostro umanesimo a disegnarne il futuro”.

Tra le voci invece aperte al progresso tecnologico, v. l’intervista al neuroscienziato, autore di basilari studi sull’empatia tra uomini, Vittorio Gallese *L’intelligenza artificiale cambierà il nostro cervello, ma evitiamo scenari apocalittici*, cit., secondo cui “Platone attaccava la poesia, la scrittura, ma già Socrate, suo maestro, se la prendeva con i libri. Poi Baudelaire era contro la fotografia, e così via, per un incessante susseguirsi di proclami nella storia, ogni volta che viene introdotta una nuova tecnologia si rimpiange quello che c’era prima. Si è detto che la fotografia avrebbe distrutto la pittura, che la televisione avrebbe cancellato il cinema. Si dice che le piattaforme distruggeranno i film, e invece non si sono mai visti tanti film come da quando ci sono le piattaforme. Demonizziamo la tecnologia, perché siamo un misto di conservazione e di rivoluzione. Quello che prevale è il cambiamento, la ricerca di nuove condizioni. C’è questa natura prometeica che ci porta a esplorare nuove frontiere, cercare nuovi orizzonti, nuove sponde, che è tipica della cultura occidentale. Siccome però ogni cambiamento presuppone un adattamento, ci sono persone che privilegiano lo status quo. E questo lo si vede anche in politica, con i conservatori o i rivoluzionari. Questa dialettica tra cambiamento e conservazione è un’altra costante della nostra specie. [...] L’esercizio più difficile oggi è quello di mantenersi in equilibrio tra un riflesso pavloviano che ci porta ad essere apocalittici, e il lodare le magnifiche sorti progressive acriticamente. L’atteggiamento più maturo è cercare di capirne di più. Studiare l’impatto che queste tecnologie stanno avendo sulle nuove generazioni, cercando di capire dove stiamo andando sulla base dei dati”. Saggiunge lo studioso che “anche l’idea che con lo sviluppo dei sistemi di intelligenza artificiale il cervello cambi non la vedo come uno scenario distopico o apocalittico, perché il cervello ha una sua natura plastica”. E aggiunge: “Sicuramente come ogni progresso tecnologico quello dell’intelligenza artificiale comporta incognite e a seconda del grado di pessimismo od ottimismo di ognuno di noi si può vedere il bicchiere mezzo vuoto o mezzo pieno. E successo con ogni tecnologia, con l’elettricità. Con l’energia nucleare, abbiamo la cobaltoterapia con cui trattiamo certe forme di neoplasie, ma c’è anche lo spettro dell’apocalisse nucleare. La metafora più calzante con l’intelligenza artificiale è quella del *pharmakon*, che dal greco è cura e veleno allo stesso tempo. L’intelligenza artificiale non è altro che una riproposizione del *pharmakon*. L’unica cosa da evitare è essere apocalittici, chiudendosi a riccio, o peggio cercare di bloccare lo sviluppo di queste tecnologie. Bisogna cercare di comprendere questa evoluzione e guidarla”.

(97) L’automazione decisoria precluderebbe quelle che Luhmann definisce “illegalità praticabili”, che in qualche caso anche il giudice (oltre alla p.a.) può essere portato a comprendere e giustificare. Secondo Luhmann un comportamento illegale è quello che lede attese formali, ma, sotto il profilo sistemico, comportamenti che si collocano in una sorta di zona grigia tra il legale e l’illegale, pur di fatto frustrando l’attesa formale, possono risultare utili (N. Luhmann, *Funktionen und Folgen formaler Organisation*, vol. V, Berlin, Duncker & Humblot, 1999, 304). Luhmann elenca diversi esempi, come il seguire regole sulla base di motivi o scopi non permessi, il comportamento secondo diritto ma adottato intempestivamente, il discostarsi consuetudinario da norme obsolete o da norme formulate in modo problematico per motivi di interesse (economico) o il discostarsi da norme in modo non significativo (futile), il comportamento contrario al diritto che al contempo evita, però, di pregiudicare altri interessi. Tutti questi comportamenti, devianti ma utili, implicano strategie adattative delle quali un membro di un’organizzazione si vale quando debba aggirare problemi di legalità, favorendo comportamenti creativi e adattamenti a un ambiente in continuo mutamento.

(98) La tesi sembra condivisa da L. Nazzicone, *op. cit.*, che stigmatizza l’eccesso di ricorsi temerari o dilatori in Cassazione che la Sezione filtro cerca di drenare e che una buona AI potrebbe aiutare.

(99) V. la giurisprudenza menzionata nel par. 3.1. Sul tema, v. C. Napoli, *op. cit.* Per altri argomenti costituzionali ostativi al giudice-robot, si rinvia alle osservazioni di L. Viola in nt. 105.

Infatti, l'uso di AI in *fase processuale* (istruttoria e soprattutto decidente), pone rilevanti problemi così schematizzabili: a) la compatibilità costituzionale di una istruttoria, di un dibattimento e di una decisione demandata ad una AI con il diritto alla difesa, con il giusto processo (artt. 24 e 111 Cost.) e con l'attribuzione costituzionale all'uomo, giudice naturale che sottostà solo alle leggi (e non ad algoritmi), del potere decisionale (artt. 102 e 101 Cost.) (100); b) pur tutelando l'ordinamento “*lo sviluppo della ricerca scientifica e tecnica*” (art. 9 Cost.), va preservata una “*riserva di umanità*” (ovvero una sfera decisoria umana) anche nell'istruire un procedimento, o nell'adottare un provvedimento del giudice o della p.a., in ossequio a canoni costituzionali (artt. 2, 3, 24, 28, 97, 98, Cost.), eurounitari (art. 6, Cedue e Regolamento europeo su AI del 2024) e nazionali (artt. 24, 101, 102, 111 Cost.; l. n. 241/1990, art. 12, d.lgs. n. 82/2005); c) vanno valutati e legislativamente allocati i rischi, anche in sede giudiziaria per la redazione di una sentenza, di danni cagionati da errori di intelligenze artificiali più o meno complesse ed autosufficienti (su programmatori, utilizzatori, manutentori, addestratori, ecc.) individuando i corretti parametri normativi (artt. 2043, 2050, 2051, 2048, 2047, 1128 c.c., art. 114 codice consumo) che si intersecano con la normativa sulla responsabilità civile del Magistrato.

La macchina intelligente non può dunque che essere ancillare supporto all'insostituibile ragionamento umano del magistrato in carne ed ossa (ma lo stesso vale per la delicata attività consultiva e difensiva di un avvocato o per l'attività consultiva ed elaborativa di un notaio), che conserva la paternità esclusiva della “sentenza” quale *prodotto cerebrale di una mente umana*, in quanto l'interprete resta e resterà sempre l'unico soggetto capace di inserire le corrette variabili date dalle argomentazioni: disattendere questo semplice auspicio, illudendosi che una macchina molto intelligente possa sostituire una mente umana saggia, competente, ricca di esperienze, di sensibilità e di amore, di spiritualità, di etica, di buon senso e di coscienza, significa aver abdicato alla *vita reale* e all'uso dell'intelligenza umana, così avviandoci, anzitempo, alla morte cerebrale per encefalogramma piatto.

È dunque a nostro avviso un errore metodologico classificare la Giustizia tra le scienze esatte a cui sarebbe agevole applicare l'AI anche per la redazione di una sentenza. Le norme giuridiche e il pensiero giuridico non conoscono infatti un'evoluzione lineare, come invece le leggi empiriche (quelle delle “scienze esatte”). La scienza e la tecnologia non dicono la verità, ma dicono cose esatte. La Giustizia deve tendere invece alla *verità*, deve tendere a coniugare la verità sostanziale con la verità processuale. E questa prerogativa valutativa è solo e soltanto dell'Uomo-magistrato!

Vale per il diritto, ma vale anche per altre “scienze non esatte”, per l'arte, la musica (101), e, soprattutto, per la vita di relazione tra esseri umani, che non può essere sostituita, se non in alcuni casi, da una IA (102).

Il problema dei problemi è dunque, e conclusivamente, quello del “*sensu del limite*” nell'utilizzo di AI, anche in sede giudiziaria: un conto è la informatizzazione e digitalizzazione della giustizia, altra cosa è la sostituzione di un giudice con una *res* intelligente.

La virata tecnologica che stiamo vivendo, se venisse esasperata, esporrebbe dunque la società, e non solo in contesti giudiziari, alla perdita di valori spirituali, quali il buon senso, la gratitudine, la solidarietà, il rispetto, la coesione sociale, alla scomparsa di professionalità cerebrale e talenti di milioni di lavoratori, alla rinuncia, in fin dei conti, alla *vita reale*, anche quella che pulsa in un complesso giudizio innanzi ad un tribunale, dove l'intelligenza algoritmica non può sostituirsi alla sensibilità umana, alla “*coscienza*” (*in primis* quella del magistrato) che,

(100) Il “diritto ad una decisione umana” viene individuato come principio processuale generale, seppur implicito, da V. Tenore, *Commentario al Codice della Giustizia Contabile*, Napoli, Esi, 2024, sub art. 2, 13 ss. Sul tema v. anche G. Gallone, *Riserva di umanità e funzioni amministrative*, Padova, Cedam, 2023. Sulla responsabilità del Magistrato v. V. Tenore, *Il Magistrato e le sue quattro responsabilità*, Milano, Giuffrè, 2016, 2 ss., tematica che andrà raccordata con la responsabilità dell'AI fallace che abbia indotto in errore il Magistrato che la abbia utilizzata.

(101) Di recente, i cantanti Giuliano Sangiorgi dei Negramaro e Jovanotti, in una diretta su Instagram, si sono fatti realizzare da ChatGPT un brano musicale in stile Negramaro. Lo segnala *online* il 7 luglio 2023 la rivista *Open.online*. Ma molto attuali sono le polemiche su programmi come LyricJam, creato dai membri del Natural Language Processing Lab dell'Università di Waterloo, in Canada. Partendo dalle esibizioni dal vivo dei musicisti, l'intelligenza artificiale riceve *clip* audio grezze, apprende le espressioni musicali e le rielabora attraverso la rete neurale. Così facendo, genera testi ritenuti idonei allo stile che recepisce.

Tra gli altri programmi di AI possono essere ricordati: AIVA, un *software* che crea musica con AI per varie attività come *jingle*, *videoclip*, *spot*, videogiochi e colonne sonore per i film; Amadeus Code, invece, è dotato di un motore AI che utilizza progressioni di accordi famosi per comporre melodie con il software iOS; Humtap è invece un programma per i musicisti che vogliono scoprire dove può portare una piccola ispirazione: basta accennare una breve melodia e ci penserà l'applicazione a generare il resto della canzone, aggiungendo beat e voce a piacimento. Poi c'è la piattaforma Brain.fm per creare musica con AI, che si basa sull'idea che la musica possa contribuire a rendere il nostro cervello più produttivo.

(102) Come ricordato in nt. 55, alcune utili sperimentazioni di macchine intelligenti utilizzate non solo come domotica, ma per far compagnia a persone anziane prive di affetti, con dialoghi tarati sull'interlocutore, ci sembrano condivisibili e da portare avanti: del resto, già da anni, una peculiare macchina (moderatamente) intelligente, ovvero la TV, è di quotidiana compagnia per milioni di anziani, sovente lasciati troppo soli da parenti assorbiti da proprie vite complesse ed impegnative. Alla solitudine esistenziale della terza età possono dunque sopperire egregiamente delle duttili macchine intelligenti, capaci di sostenere, con maggiore pazienza e tolleranza di molti esseri umani, dialoghi e ascolto di racconti anche ripetitivi di anziani. Analoghi rilievi valgono per programmi interattivi “di compagnia” (anche sessuale) per uomini o donne soli o problematicamente curiosi.

secondo l'insegnamento dantesco, è quella parte di noi “che vive e sente e sé in sé rigira” (*Purgatorio*, XXV canto) e che ci permette di percepire e comprendere il significato della realtà fisica, delle azioni umane, delle emozioni e dei pensieri senza una mera e cieca traduzione di segnali in altri segnali della stessa natura, come una mera AI (103).

Come ha ben colto Erik Larson nel suo *Il mito dell'Intelligenza artificiale*, l'intelligenza artificiale induttiva continuerà a fare passi avanti per quanto riguarda l'esecuzione di compiti limitati, ma se desideriamo un progresso reale, dovremo iniziare ad apprezzare appieno la sola vera intelligenza che conosciamo: la nostra.

Tali approdi, fondati su una visione laica della vita, trovano conforto anche in un'ottica cristiana: difatti il Papa è consapevole del problema, e non casualmente ha dedicato la *Giornata della pace* del 1° gennaio 2024 all'intelligenza artificiale e alle sue sfide. Come ben colto da attenti studiosi di matrice cattolica (104), se l'intelligenza artificiale elaborerà pensieri, se la robotica li tradurrà in opere, costruzioni e produzioni, se potremo riprodurci in laboratorio, se alle domande sul senso della vita risponderà il Dr. Google, che cosa rimarrà dell'uomo nell'uomo? Cosa ne sarà della sua coscienza, del suo Spirito del suo ingegno, della sua creatività? Saremo ancora a immagine e somiglianza di Dio o di un algoritmo? Sono gli interrogativi che, secondo tale dottrina, spingono a dedicare al tema dell'intelligenza artificiale (IA) e delle sue sfide la prossima Giornata della pace. Nuove preoccupazioni si aggiungono a quelle della crisi climatica che segnala il rischio di una rottura dolorosa dell'equilibrio uomo-natura.

Oggi, conclude tale scritto, è in atto una rivoluzione che sta toccando i gangli essenziali dell'esistenza umana. Stanno mutando le modalità essenziali di intendere il generare, il nascere e il morire. È messa in discussione la specificità dell'essere umano rispetto al creato, la sua unicità nei confronti degli animali e la sua relazione con macchine e intelligenza artificiale. Papa Francesco sostiene pertanto che nel mondo liquido di oggi c'è bisogno di un nuovo umanesimo per “*ripensare alla presenza dell'essere umano nel mondo*”.

Dunque, la vera intelligenza, anche nel giudicare in un processo, non è algoritmica, ma è nella capacità di comprendere, di *intus-legere* (leggere dentro), di contestualizzare, di capire in profondità e di trovare connessioni insospettite. Come ha ben sintetizzato uno dei massimi fisici italiani (105), “la vera intelligenza è intuizione, immaginazione, reattività, ingegno e inventiva. È lungimiranza, visione e saggezza. È empatia, compassione, etica e amore. È integrazione di mente, di cuore e di azioni coraggiose. In altre parole, la vera intelligenza non è separabile dalle altre proprietà che ci rendono umani e che richiedono la coscienza ed il libero arbitrio, cioè la capacità di comprendere e quella di prendere decisioni inaspettate, creative ed etiche. Le macchine non potranno mai fare queste cose perché, se fossero libere come siamo noi, sarebbero più pericolose che utili. Esse funzionano, ma non capiscono. E capire non è riconducibile ad un algoritmo”.

(103) Come rimarca F. Faggin, *op. cit.*, 126, “quando la coscienza è coinvolta nell'apprendimento, essa ci permette di raggiungere livelli di competenza nella percezione, nel riconoscimento e nella comprensione che eccedono incommensurabilmente le capacità delle macchine. [...] Quando si parla di intelligenza, è la comprensione cosciente che fa la differenza! E la comprensione è una proprietà esclusiva, non algoritmica, della coscienza”. L'autorevole studioso conclude dunque (p. 139) che il “riconoscimento cosciente va ben oltre le capacità delle macchine, che si limitano al riconoscimento di pattern simbolici che non richiedono comprensione. Tale apprendimento automatico basato su reti neurali artificiali è infatti inconscio. Le macchine possono imparare, ma non percepiscono e non comprendono nulla perché non sono coscienti”.

Anche M. Chiriatti, *Incoscienza artificiale*, Roma, Luiss University Press, 2021, osserva come “la coscienza è incomputabile” e che sarebbe meglio, come evidenzia il titolo del suo volume, sostituire la parola “intelligenza”, che ha un'accezione positiva, con “incoscienza”, visto che gli algoritmi, eseguendo regole che imparano autonomamente dai dati, “producono risultati senza alcuna comprensione e coscienza di ciò che stanno facendo”.

(104) Le riflessioni riportate sono di A. Mattioli, *A immagine e somiglianza di Dio o dell'algoritmo. Il cristiano nell'era dell'intelligenza artificiale*, in <www.famigliacristiana.it>, 26 agosto 2023.

(105) F. Faggin, *op. cit.*, 143. Ma già P.L.M. Lucatuorto, S. Bianchini, *Discrezionalità e temperamento degli interessi nei processi decisionali dall'amministrazione digitale*, in *Cyberspazio e dir.*, 2009, 43, rilevavano giustamente come “i sistemi per la decisione automatica [...] [si limitino] a cogliere solo l'aspetto più superficiale della conoscenza e del ragionamento giuridico”. Osserva anche L. Viola, *L'intelligenza artificiale nel procedimento e nel processo amministrativo*, cit., 8 ss. in relazione alla adozione di atti discrezionali della p.a. (ma le conclusioni valgono a maggior ragione per il ragionamento di un giudice in una sentenza), «Con tutta evidenza, gli ampi riferimenti ai concetti giuridici precisi e determinati, a forme di ragionamento automatiche o procedurali ed il parallelo con il sillogismo giuridico evidenziano un contesto complessivo in cui risulta sicuramente suscettibile di automazione l'emanazione degli atti vincolati, ovvero delle sole fattispecie in cui il processo di decisione si fonda sull'“applicazione di regole” e non sull'esercizio della discrezionalità amministrativa».

Lo stesso A., in un più recente scritto (L. Viola, *La giustizia predittiva del lavoro*, cit., 12 ss.) nega la possibilità di un giudice-robot ritenendo che “La diffusione di un giudice robot, sostitutivo del giudice, non è applicabile in Italia, almeno per le seguenti tre convergenti ragioni: l'art. 25 Cost., nel predicare il principio di giudice naturale precostituito per legge, ne postula la sua umanità; l'art. 51 c.p.c. impone l'astensione del giudice per grave inimicizia con una delle parti; ebbene: l'inimicizia è concetto tipicamente umano, con la conseguenza che il legislatore postula una persona umana, non robot; sussiste un principio di simmetria di natura tra giudicante e persona giudicata, nelle principali declinazioni del processo civile, penale, amministrativo; si intende dire che il giudicante deve essere umano come il soggetto che va a giudicare per la ragione decisiva che, diversamente, non riuscirebbe a comprendere le pulsioni fino in fondo, che sono tipicamente umane; la conseguenza sarebbe una limitazione alla difesa, in contrasto con il suo carattere di inviolabilità ex art. 24 Cost., in quanto il giudice non riuscirebbe a comprendere fino in fondo la difesa perché di natura diversa”.

Sul tema, v. L. Cappello et al., *op. cit.*

Tali chiarissime conclusioni calzano, come un capo sartoriale, alla intelligente e complessa decisione di un magistrato, non delegabile a una “semplice” macchina.

* * *

IL D.D.L. (N. 1066 AS), DI INIZIATIVA GOVERNATIVA, APPROVATO IL 23 APRILE 2024, SULLA INTELLIGENZA ARTIFICIALE. I PRINCIPI FONDAMENTALI DELL'AI PER IL SUO UTILIZZO IN ITALIA, IN LINEA CON "L'AI ACT" DELIBERATO DALL'UE. LE DISPOSIZIONI DI SETTORE. LA STRATEGIA NAZIONALE E LE NUOVE AGENZIE ISTITUITE. LA TUTELA PER GLI UTENTI E PER IL DIRITTO D'AUTORE E LE PENE PER I TRASGRESSORI. IL G7 SVOLTOSI IL 13-15 GIUGNO 2024 A BORGO EGNAZIA (BR), CON L'INTERVENTO DI PAPA BERGOGLIO

di Maria Teresa D'Urso (*)

Abstract: Il presente lavoro analizza l'iniziativa assunta dal Presidente del Consiglio dei Ministri, che il 23 aprile 2024 ha approvato il d.d.l. (n. 1066 AS) per l'introduzione di disposizioni in materia di IA. Si esaminano *in primis* i principi fondamentali per l'utilizzo in Italia di questa straordinaria tecnologia, fornendo soluzioni equilibrate per la gestione del rischio, fondate su una visione antropocentrica, che pone la volontà umana al centro del processo decisionale, lasciando all'AI, una funzione servente e di supporto, subordinata alla decisione finale dell'uomo, che è anche etica e sociale. Vengono poi analizzate le "disposizioni di settore", meritevoli di maggiore attenzione quali: la sanità e disabilità; il lavoro; le professioni intellettuali; la pubblica amministrazione; le attività giudiziarie e la cybersicurezza nazionale. Si prendono inoltre in esame i cinque ambiti importanti nel progetto del Governo: la strategia nazionale messa in campo, le due Autorità nazionali nominate: l'Agenzia per l'Italia Digitale (AgID) e l'Agenzia per la cybernetica nazionale (Acn), con il fine di incentivare, controllare e garantire il corretto utilizzo dell'IA secondo i principi fissati dal presente d.d.l. e dall'Ue. Vengono infine valutate le azioni di promozione messe in campo, supportate con il finanziamento di un miliardo di euro, reso disponibile con l'intervento di "Cdp Venture Capital", s.p.a., società del Ministero delle imprese e del made in Italy (MimIt), che investe attraverso *start-up* e Pmi in venture capital. Vengono poi analizzate le tutele previste a tutela degli utenti e dei titolari del diritto d'autore, nonché le sanzioni penali per i trasgressori. Infine, si sono esaminati i risultati della sessione del G7 svoltosi il 13-15 giugno 2024 a Borgo Egnazia di Fasano (BR), con l'intervento di Papa Bergoglio Francesco e dei grandi della terra, avente ad oggetto anche il ruolo dell'IA nel mondo.

This work analyzes the initiative taken by the Council of Ministers, which on 23 April 2024 approved the bill for the introduction of provisions regarding artificial intelligence. We first examined the rules of principle and sector provisions for the use of this new, extraordinary technology in Italy, providing balanced solutions for risk management based on an anthropocentric vision, which places human will at the center of the decision-making process, leaving AI with a serving and support function, however subordinate to man's final decision, which is also ethical and social. The sector provisions worthy of greater attention are then analysed, such as: healthcare and disability; Work; intellectual professions; public administration; judicial activities and national cybersecurity. Furthermore, the five important areas for the Government's project are examined: the national strategy implemented, the national authorities appointed, namely the Agency for Digital Italy (AgID) and the Agency for National Cybernetics (ACN), in order to incentivize, control and guarantee the correct use of artificial intelligence according to the new law and the principles established by the EU. Finally, the promotional actions implemented are evaluated, supported with the financing of one billion euros, made available with the intervention of Cdp Venture Capital, s.p.a., a company of the Ministry of Business and Made in Italy (MIMIT), which invests through startups and SME in venture capital. We then analyzed the protections provided for the defense of users and copyright holders, and the criminal sanctions for violators. Finally, the results of the G7 held on 13-15 June 2024 in Borgo Egnazia di Fasano (BR) were analysed, with the intervention of Pope Bergoglio Francis and the world's greats, concerning the role of AI in the world.

Sommario: 1. Premessa. – 2. I principi fondamentali per l'utilizzo in Italia dell'IA nei settori produttivi. – 3. Le Disposizioni per i Settori tutelati: sanità e disabilità; lavoro; professioni intellettuali; pubblica amministrazione; attività giudiziaria; cybersicurezza nazionale. – 4. La strategia nazionale e le nuove Agenzie: AgID e Acn ed i finanziamenti deliberati. – 5. Tutela per gli utenti italiani e in materia di diritto d'autore. – 6. Tutela penale. – 7. Il G 7 svolto il 13-15 giugno 2024 a Borgo Egnazia di Fasano (BR), con l'intervento di Papa Bergoglio, che espone il "Rome call for AI Ethics" ed indica il ruolo etico dell'IA nel mondo. – 8. Conclusioni.

(*) M.T. D'Urso è Consigliere della Corte dei conti.

1. Premessa.

Il Governo italiano, in concomitanza con l'adozione nel marzo 2024 dell'“AI Act” da parte del Parlamento Ue, rettificato il 21 aprile 2024 dal Consiglio Ue – prima ancora della pubblicazione sulla G.U. della Ue – ha approvato in data 23 aprile 2024, su proposta del Presidente del Consiglio, Giorgia Meloni, e del Ministro della giustizia, Carlo Nordio, un disegno di legge (d.d.l. n. 1066 AS), recante “*Norme per lo sviluppo e l'adozione di tecnologie di intelligenza artificiale*” (1), costituito da n. 25 articoli, distribuiti in cinque capitoli, che disciplina l'utilizzo dell'intelligenza artificiale, in una visione proattiva e cautelativa, del suo impiego in Italia.

Si precisa fin d'ora che il d.d.l. (n. 1066 AS) non si sovrappone al Regolamento Europeo approvato il 13 marzo 2024 sull'*Artificiale Intelligence* (2), ma accompagna il quadro regolatorio di tale innovativa attività negli spazi specifici del diritto interno in Italia, in particolare nel lavoro, pubblica amministrazione, giustizia, sanità, cybersecurity, produttività, sicurezza, tutela dei dati, con l'obiettivo primario di promuovere “*un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a coglierne le opportunità*”, garantendo nel contempo “*la vigilanza sui rischi economici e sociali e sull'impatto dei diritti fondamentali dell'intelligenza artificiale*”, in linea con i principi dell'Unione europea (art. 1 d.d.l.).

(1) In dottrina: L. Di Giacomo, *Intelligenza Artificiale: approvato in CdM il disegno di legge italiano* (29 aprile 2024), in www.diritto.it; F. Meta, *Intelligenza artificiale, ecco la legge italiana: spinta alla produttività, carcere per le “manipolazioni”*, in www.corrierecomunicazioni.it (23 aprile 2024); Redazione CeSDA – Centro Studi di Diritto Amministrativo, *In Parlamento il disegno di legge del Governo sull'Intelligenza Artificiale* (3 maggio 2024), in www.giurdanella.it; L. Zorloni, *Cosa ha deciso il Governo sull'Intelligenza Artificiale*, in www.wired.it; M. Borgobello e L. Daffarra, *D.d.l. Intelligenza Artificiale: così l'Italia tutela la dignità delle persone* (24 aprile 2024), in www.agendadigitale.eu; La Redazione, *Approvato dal Consiglio dei ministri un disegno di legge in materia di intelligenza artificiale* (24 aprile 2024), in www.giurisprudenzapenale.com; C. Ponti, *D.d.l. sull'intelligenza artificiale, tutela la privacy e la cybersecurity della nuova legge* (26 aprile 2024), in www.cybersecurity.it; C. Sorrentino, *Intelligenza artificiale: cosa prevede la normativa italiana* (3 maggio 2024), in <https://blog.osservatori.net>; M. Biasi, *Il lavoro nel disegno di legge italiano sull'IA: antipasto dell'IA Act* (7 maggio 2024), in www.agendadigitale.eu; Direzione Assiv, *Disposizioni e delega al Governo in materia di intelligenza artificiale* (disegno di legge), in www.assiv.it.

(2) Si ricorda che il Regolamento approvato dal Parlamento europeo in data 13 marzo 2024 è stato preceduto dal progetto di AI Act, definito nel 2023, avente ad oggetto le regole per l'utilizzo dell'IA valide per i Paesi dell'Ue nel rispetto dei diritti e dei valori dell'Unione, da sottoporre poi all'approvazione della Commissione Europea nel 2025, per entrare in vigore entro il 2026. Il Parlamento Europeo, dopo aver raggiunto l'accordo con tutti gli Stati membri, ed aver acquisito il parere degli stessi nel 2023, ha approvato il 13 marzo 2024, a larga maggioranza (voti favorevoli n. 523, contrari n. 46 ed astensioni n. 49), il Regolamento sull'Artificiale Intelligence, la prima legge al mondo, la cui finalità è quella di garantire il rispetto dei diritti fondamentali dell'Ue, la sicurezza dei cittadini, la democrazia e lo Stato di diritto, nonché la sostenibilità ambientale da parte dei sistemi di Intelligenza Artificiale c.d. “ad alto rischio”.

Tale legge era urgente ed indispensabile, dato che la materia dell'intelligenza artificiale, in particolare l'AI di “tipo generativo”, quali Chat Gtp (acronimo di *Chat Generative pretrained transformer*) di Open Ai, società privata Usa, è talmente dirompente che necessita di una regolamentazione specifica, non essendo più sufficiente il Regolamento Gdpr (acronimo di *General data protection regulation*) n. 679/2016 in materia di privacy, approvato dall'Ue ed entrato in vigore il 25 maggio 2018, relativo alla protezione dei dati personali e sensibili.

Infatti, Chat Gtp, che utilizza le tecniche di *Machine Learning* e *Dreep Learning*, è in grado di generare autonomamente immagini, musica, poesie, testi scritti che non esistevano prima, altamente realistici, che imitano la creatività stessa dell'uomo. È un dato di fatto che Chat Gtp sta rivoluzionando il modo di creare musica, migliorare la diagnostica medica, scoprire nuovi farmaci, nuove misure di sicurezza informatica, contrastare frodi, evitare evasioni fiscali, creare programmi per prevedere possibili incendi o alluvioni, decifrare in archeologia papiri carbonizzati (ritrovati ad Ercolano a seguito dell'eruzione del Vesuvio del '79 d.c.), dimostrando in tal modo di essere una preziosa risorsa in una varietà di ipotesi e settori (esempio medicina, difesa militare per utilizzo di droni, ricerche spaziali, ecc.). Ma rappresenta anche una tecnica molto pericolosa per la possibile divulgazione di “pregiudizi”, la disinformazione, i *dreepfake* ed altri contenuti non veritieri, con ripercussioni molto gravi e negative in ambito politico, commerciale, economico e per l'uso improprio di dati personali, per influenzare le elezioni di Stati rivali, ecc.

Il Regolamento del 13 marzo 2024 rappresenta un quadro giuridico omogeneo, con cui viene regolato lo sviluppo, la diffusione commerciale e l'utilizzo dei sistemi algoritmici di IA, compresi quelli di tipo generativo (esempio Chat Gtp di Open Ai, oppure Dell Be o Bard di Google, ecc.), nel rispetto delle regole e dei valori dell'Unione europea, al fine di ottenere chiarezze in un mondo opaco ed in rapida evoluzione, quale quello degli algoritmi automatizzati usati dall'IA. Risulta che, oltre agli USA, anche Cina, Russia ed India investono molto nel settore. Si precisa infine che con detto Regolamento l'Unione Europea ha concesso agli Stati membri 24 (ventiquattro) mesi di tempo perché le aziende interessate nel settore, specie quelle che producono software di intelligenza artificiale, possano adeguarsi alle nuove regole e rafforzare il copyright (“diritto esclusivo di uso di un'opera propria”), dando maggiore trasparenza sui contenuti generati dagli algoritmi (c.d. “AI Pact”), prima che la legge medesima diventi operativa.

In dottrina sull'AI Act approvato il 13 marzo 2024 dal Parlamento Europeo: B. Calderoni, *AI Act, il punto sui risultati raggiunti e i dubbi sul futuro* (12 dicembre 2023), in www.agendadigitale.eu; M. Pastiglia, *AI Act, l'intelligenza artificiale rispetti l'umano: i punti salienti del regolamento Ue* (14 giugno 2023), in www.cybersecurity360.it; M. Toti, *AI Act: con il voto del Parlamento l'Ue traccia il futuro dell'Intelligenza Artificiale* (23 giugno 2023), in www.altalex.com; M. Flora, *AI Act: tutto quello che dovete sapere* (9 dicembre 2023), in <https://linkedin.com>; A. Signorini, *L'A.I. Act è a un passo. Noi siamo pronti?* (22 maggio 2023), in www.insinitinere.it; Wired Italia, *Con l'A.I. Act l'Europa approva il primo decreto al mondo sull'intelligenza artificiale* (9 dicembre 2023), in www.wired.it.

In tal modo si intendono tracciare regole nazionali per la ricerca, lo sviluppo e la corretta applicazione di questa straordinaria tecnologia, basata su modelli digitali e su algoritmi, ormai applicata nei più diversi settori dell'attività umana.

2. I principi fondamentali per l'utilizzo dell'IA nei settori produttivi.

Come indicato in precedenza, il Consiglio dei Ministri, su proposta del Presidente del Consiglio e del Ministro della giustizia, ha approvato il 23 aprile 2024 – con la previsione della richiesta alle Camere del Parlamento di una sollecita calendarizzazione nel rispetto del regolamento dei due rami del Parlamento – il d.d.l. (n. 1066 AS) per la introduzione in Italia di disposizioni e la connessa delega al Governo in materia di intelligenza artificiale, in una visione proattiva e cautelativa del suo sviluppo ed utilizzo in Italia.

Tale iniziativa intende collocare l'Italia in una posizione primaria nel dibattito globale sull'argomento, onde assicurare che l'AI venga utilizzata secondo principi di correttezza, trasparenza e responsabilità, ossia “in una dimensione antropocentrica”, che mantenga l'uomo al centro del processo decisionale, garantendo che la tecnologia e gli algoritmi dell'AI siano soltanto di supporto ed ausilio all'attività dell'uomo, ma senza sostituirsi ad esso deputato ad assumere le decisioni finali, aventi anche connotati di natura etica e sociale.

Si prevedono, a tutela della dimensione antropocentrica, idonei strumenti di “vigilanza” per evitare potenziali rischi, sociali ed economici, che un uso improprio o un impiego dannoso di AI possa arrecare ai diritti fondamentali dei cittadini, tutelati in modo analogo negli antri Stati dell'Unione europea. In tale ottica il d.d.l. (n. 1066 AS) non si sovrappone al Regolamento Europeo sull'AI approvato il 13 marzo 2024, ma rappresenta una autonoma regolazione, negli spazi propri del nostro diritto interno, con cui si pongono le basi etiche ed operative per lo sviluppo e l'applicazione dell'AI in Italia, in linea con i principi già fissati dell'IA Act.

Il disegno di legge si sviluppa in 25 articoli, distribuiti in cinque ambiti. Molto importanti sono i “principi fondamentali”, compresi nei primi articoli (artt. 1-6 d.d.l.), con essi cui si stabiliscono le basi etiche, sociali ed operative per lo sviluppo e l'utilizzo dell'AI in Italia. In tale ambito viene delineato un quadro regolatorio che pone in primo piano la dignità dell'uomo, la sicurezza e la trasparenza nella governance degli algoritmi di AI a livello nazionale, in linea con i principi dell'AI Act già tracciati dall'Unione europea. Quindi successive norme prevedono che i sistemi ed i modelli di IA nell'ordinamento italiano ed in quello europeo devono fondarsi – oltre che sul rispetto dei diritti fondamentali e delle libertà costituzionali – anche sui principi di trasparenza, proporzionalità, sicurezza, valorizzazione anche economica del dato, protezione dei dati personali, riservatezza, non discriminazione, parità dei sensi e sostenibilità.

Viene garantita, come valore primario, l'autonomia ed il potere decisionale dell'uomo, nel senso che l'AI deve servire come strumento tecnico per potenziare, sul piano servente e subordinato, il potere decisionale dell'uomo, che resta insostituibile in quanto valuta anche tenuto conto di elementi sociali, etici, morali, di giustizia sostanziale, garantendosi sempre e comunque che le c.d. “decisioni automatizzate” possano essere scrutinate, impugnate, modificate o riformate dalla decisione finale dell'uomo. Si garantisce poi alle persone con disabilità il pieno accesso ai sistemi di intelligenza artificiale, senza alcuna forma di discriminazione.

È stabilita poi la regola che l'utilizzo dell'IA non deve avere una influenza distorsiva sulla vita democratica del Paese, protetta attraverso una vigilanza particolare contro l'uso di tecnologie, come i *deepfake*, capaci di manipolare l'opinione pubblica ed interferire con il libero svolgimento delle elezioni democraticamente svolte. In tale contesto, la legge individua una tutela rafforzata in tema di cybersecurity contro tale rischio, che riguarda tutto il ciclo di vita dei sistemi e dei modelli di intelligenza artificiale.

Nei settori produttivi viene garantito lo sviluppo economico e l'accesso ai dati. Infatti, si prevede la promozione di un ecosistema dell'AI che sia innovativo, equo e competitivo, incentivando la creazione di un mercato dell'AI attraverso appalti pubblici specifici, facilitando l'accesso a dati di alta qualità per le aziende (pubbliche e private) e la comunità scientifica, con il fine di colmare il divario tra pubblico e privato, che sussiste in questo specifico settore tecnologico. Ciò con il dichiarato proposito di migliorare la produttività nazionale ed avviare nuove attività economiche che migliorino il benessere sociale, nel rispetto del principio di libera concorrenza nel mercato dell'utilizzo di dati di alta qualità. È quindi previsto che le piattaforme di *e-procurement* (acronimo di *Electronic Procurement*) delle pubbliche amministrazioni seguano le medesime regole.

È anche stabilito, quale principio primario della legge, che l'utilizzo dei sistemi di IA nei mezzi di comunicazione venga realizzato senza pregiudizio per i principi di libertà e di pluralismo di espressione, e per il diritto alla obiettività, imparzialità e lealtà dell'informazione. Da ultimo, ma non per importanza, viene affrontato anche la questione del consenso per l'utilizzo dei sistemi di AI da parte dei minori, proponendo – in linea con analoghi criteri già adottati da altri Paesi dell'Unione europea – soglie di età che sono differenziate per l'accesso a tecnologie basate sull'AI, e per il consenso dei dati personali (art. 4 d.d.l.). Tale ultima disposizione rappresenta una oggettiva e reale preoccupazione per la protezione della privacy e per ragioni etiche nell'uso della intelligenza artificiale, dato che è necessario un approccio molto cauto ed informato nei confronti di tecnologie potenzialmente invasive e molto libere nei contenuti.

3. Le Disposizioni per i Settori tutelati: Sanità e disabilità; lavoro; professioni intellettuali; pubblica amministrazione; attività giudiziaria; cybersicurezza nazionale.

Il disegno di legge in esame prosegue, poi, con specifiche disposizioni che riguardano in particolare il settore della sanità, il mondo del lavoro, l'esercizio delle professioni intellettuali, l'attività della pubblica amministrazione, lo svolgimento della funzione giudiziaria e la realizzazione di una cybersicurezza nazionale, efficiente e professionale, a protezione dello spazio cibernetico, ossia dei sistemi, reti e programmi di obiettivi sensibili (cfr. ministeri, società pubbliche o private, banche, siti militari, ecc.) da parte di attacchi digitali di hacker anche stranieri.

- Sanità e disabilità (cfr. artt. 7-9 d.d.l.). In tale settore è previsto che l'utilizzo dell'intelligenza artificiale non può in alcun modo essere usata per selezionare i pazienti con criteri discriminatori, ponendo condizioni e/o restrizioni limitative, in via diretta o indiretta, per l'accesso alle prestazioni sanitarie, che costituiscono un diritto garantito a tutti i cittadini (art. 32 Cost.), la cui lesione comporta anche il risarcimento del danno sofferto. Prioritario poi nel settore è il diritto dell'interessato ad essere informato circa l'utilizzo dei sistemi di IA. In ambito sanitario deve comunque lasciarsi impregiudicata la spettanza della decisione alla professione medica, essendo l'uso della macchina solo un supporto di ausilio per la decisione finale riservata al personale sanitario.

È prevista poi la messa in funzione di una "piattaforma di IA nazionale", con il fine di contribuire ad una migliore offerta di cure sanitarie, in particolare per l'assistenza territoriale, la cui progettazione è affidata all'Agens (acronimo di "Agenzia nazionale per i servizi sanitari regionali"), braccio operativo del Ministero della salute nel digitale. Detta Agenzia dovrà affiancare i medici di base per la presa in carica, fornire consigli clinici e consentire ai cittadini di accedere alle "Case di comunità", che è una rete di servizi sanitari locali in via di sviluppo, in particolar in Lombardia e nel resto del Nord. È stabilito poi – nell'ambito della ricerca e della sperimentazione scientifica con sistemi di IA in ambito sanitario – che i trattamenti dei dati, anche personali, eseguiti da enti pubblici o privati, senza scopo di lucro per la ricerca e la sperimentazione scientifica "siano dichiarati di rilevante interesse pubblico".

Si consente in tal modo agli enti *no-profit* la legittimazione all'uso secondario di dati personali, anche di categorie particolari, senza che sia necessario richiedere agli interessati un ulteriore consenso. In tal modo si impone ai centri studio, attraverso i loro comitati etici, di comunicare al Garante della privacy l'avvio dei loro studi, senza ottenere una autorizzazione preventiva. Sarà poi l'ente, se lo ritiene, a richiedere informazioni sul dossier. Il perimetro della norma come innanzi formulata si applica anche a ricerche su medicinali, nuove cure, ma anche a sensori, protesi ed interfacce uomo-macchina.

- Lavoro (cfr. artt. 10-11 d.d.l.). In questo settore si richiamano principi già fissati nello Statuto dei lavoratori (l. 20 maggio 1970, n. 300, novellato dalla l. 23 dicembre 2021, n. 238 e dalla sentenza Corte cost. 19 maggio 2022, n. 125), e nel Gdpr (acronimo di *General data protection regulation*), che è il Regolamento n. 2016/679 approvato dall'Unione europea il 27 aprile 2016 ed entrato in vigore il 24 maggio 2018, in materia di trattamento dei dati personali e di *privacy*.

In particolare, si pongono regole a tutela del lavoratore contro il rischio di possibili abusi, o usi impropri, o dannosi dei sistemi di Intelligenza Artificiale, vietando l'impiego di tali strumenti quando essi abbiano finalità discriminatorie. Si applica al mondo del lavoro il principio antropocentrico dell'utilizzo di IA, nel senso che gli algoritmi di IA possono utilizzarsi soltanto per migliorare le condizioni di lavoro, tutelare l'integrità psico-fisica del lavoratore, accrescere le qualità delle prestazioni lavorative e la produttività delle persone, in linea con le stesse regole del diritto praticato dai Paesi della Unione europea. Il principio è che *"l'utilizzo dell'intelligenza artificiale in ambito lavorativo deve essere sicuro, affidabile, trasparente e non può svolgersi in contrasto con la dignità umana né violare la riservatezza dei dati personale"*. Quindi qualora il datore di lavoro, pubblico o privato, faccia ricorso a sistemi di AI, è obbligato necessariamente ad informarne i dipendenti attraverso idonei mezzi di comunicazione e, nel contempo, assicurarsi che siano rispettati i diritti del settore.

Inoltre, il d.d.l. (n. 1066 AS) ribadisce il principio di equità e di non discriminazione, stabilendo in particolare che l'utilizzo dei sistemi di IA non può determinare discriminazioni di alcun genere nell'organizzazione o nella gestione del rapporto di lavoro. A tale riguardo viene istituito presso il Ministero del lavoro e delle politiche sociali un Osservatorio sull'adozione di sistemi di IA nel mondo del lavoro, per il monitoraggio e la verifica della costante applicazione di tale principio.

È anche prevista *"la possibilità di un riconoscimento di un equo compenso modulabile sulla base della responsabilità e dei rischi connessi all'uso dei sistemi di intelligenza artificiale"*. Tale circostanza deve però essere ulteriormente precisata, in termini giuridici e fattuali, sulla base perimetrica dei principi vigenti in tema di indennizzo per attività connesse al mondo del lavoro, all'impegno concreto che comporta l'utilizzo della AI nei singoli luoghi di lavoro.

- Professioni intellettuali (cfr. art. 12 d.d.l.). Si stabilisce poi in tale settore la regola – esplicativa del principio antropocentrico – che *"il pensiero critico dell'uomo"* debba essere sempre prevalente rispetto all'utilizzo degli strumenti dell'intelligenza artificiale, che si pongono in funzione strumentale e di supporto rispetto all'attività dell'uomo, la quale resta prevalente e decisiva per la soluzione finale da assumere. Quindi l'uso di algoritmi si

qualifica come attività ausiliaria rispetto a quella del professionista e non quale sostitutiva di essa. A tale riguardo è stabilito che le informazioni riguardanti i sistemi di intelligenza artificiale utilizzati dal professionista, nell'espletamento del singolo incarico (giuridico, tecnico, contabile, ecc.) siano comunicate al cliente in modo chiaro, semplice e completo. In particolare, deve farsene espressa menzione nell'accordo scritto che viene concluso all'atto del conferimento dell'incarico. Tali regole si ritiene debbano tener conto e sono da bilanciare con la tutela del *know how* e dei segreti commerciali di ciascun professionista.

Infine, si evidenzia, sul piano delle professioni intellettuali, che un aspetto di particolare interesse è rappresentato dall'introduzione dell'intelligenza artificiale nelle tecniche dei "controlli fiscali", con il dichiarato obiettivo di dare un significativo aiuto collaborativo all'Agenzia delle entrate, con il fine dichiarato di semplificare le procedure per i contribuenti e migliorare l'efficienza dell'amministrazione fiscale e per combattere l'annoso problema dell'evasione fiscale.

- Pubblica amministrazione (cfr. art. 13 d.d.l.). Viene posta la regola primaria, la quale insegna che l'utilizzo della IA deve garantire il buon andamento e l'efficienza dell'attività amministrativa, ridurre i tempi di definizione dei procedimenti ed aumentare le qualità e la quantità dei servizi erogati ai cittadini ed alle imprese.

Il compito di introdurre l'utilizzo dell'intelligenza artificiale negli enti pubblici viene affidato all'Agenzia per l'Italia digitale (Agid), di nuova istituzione, che ha anche la funzione di certificare ed accreditare delle società che sviluppino gli algoritmi. Inoltre, l'Agid è in collegamento funzionale con la Consip (acronimo di Concessionaria dei Servizi informatici pubblici), società costituita in modalità di s.p.a., di proprietà del Ministero dell'economia e delle finanze, la quale funge da tempo, com'è noto, da centrale acquisti della pubblica amministrazione ed è abilitata a ricevere la certificazione di qualità ISO 9001:2008 per i processi d'acquisto di beni e servizi. In concreto l'Agid ha il compito precipuo di fornire alla Consip gli indirizzi su programmi ed aziende che dovranno impiantare ed utilizzare l'AI nei propri programmi operativi.

Nello specifico si auspica che *“vengano privilegiate quelle soluzioni che garantiscono la localizzazione e la elaborazione dei dati critici presso <data center post> sul territorio nazionale. Nonché modelli fondativi in grado di assicurare elevati standard in termini di trasparenza sulle modalità di addestramento e di sviluppo di applicazioni basate su AI generativa”*. Un ruolo importante a tale riguardo viene svolto dal Polo strategico nazionale (Psn), società che è l'infrastruttura ad alta affidabilità, la quale ha l'obiettivo di dotare la pubblica amministrazione di tecnologie e infrastrutture *cloud*, che possono beneficiare delle più alte garanzie di affidabilità, resilienza ed indipendenza.

Il Dipartimento per la trasformazione digitale – com'è noto – ha realizzato il Polo strategico nazionale (Psn) attraverso una convenzione con una società di nuova costituzione, la Psn, s.p.a., in cui sono raccolti tutti i dati ed i servizi critici e strategici delle amministrazioni centrali (circa 200), delle Aziende sanitarie locali e delle principali amministrazioni locali (Regioni, Città Metropolitane e Comuni con oltre 250 mila abitanti). Quindi l'ingresso dell'AI negli enti pubblici avverrà attraverso la guida professionale di Psn, con l'intesa che l'AI nel settore pubblico dovrà essere improntata al principio di trasparenza, *“assicurando agli interessati la conoscibilità del suo funzionamento e la tracciabilità del suo utilizzo”*, ed essere usata a supporto di decisioni assunte da personale umano, responsabile della propria funzione, in ossequio al principio antropocentrico, che informa tutta la normativa in tema di AI. L'obiettivo 2026 che intende raggiungere Psn è che almeno il 75 per cento delle pubbliche amministrazioni italiane possano utilizzare servizi in *cloud*, quale obiettivo inserito nel Pnrr per l'Italia.

L'utilizzo di AI nel settore pubblico deve assicurare – come già accennato – la conoscibilità e la tracciabilità del suo utilizzo, e svolgere solo una funzione strumentale e di supporto, cioè servente e strumentale rispetto al potere decisionale della persona umana, dotata di responsabilità ed autodeterminazione, secondo le regole generali fissate dall'art. 28 Cost, essendo solo l'uomo che deve decidere, in adesione anche nel settore pubblico del principio antropocentrico cui si ispira il presente d.d.l. (n. 1066 AS).

Attività giudiziaria (cfr. art. 14 d.d.l.). Si ribadisce che nell'amministrazione della giustizia l'utilizzo dei sistemi di intelligenza artificiale è consentito solo per finalità strumentali e di supporto, utili per agevolare l'organizzazione e la semplificazione del lavoro giudiziario e la ricerca di precedenti giurisprudenziali o di articoli di dottrina, agevolando lo studio e la soluzione di casi analoghi, quando vi sia da definire una controversia o una vertenza simile.

In osservanza del quadro costituzionale (artt. 102, 103, 111, c. 1, Cost.), possono individuarsi attività che il giudice non può delegare agli algoritmi, quali la *“interpretazione letterale”* (c.d. *vox iuris*), la *“interpretazione logica”* (che, superando il significato immediato della disposizione, mira a stabilire lo scopo che il legislatore ha inteso realizzare, emanandola), e la *“analogia legis”* (ammissibile solo se basata su determinati presupposti normativamente prefissati). Si tratta di funzioni riservate in via esclusiva al giudice per la soluzione del caso concreto, il quale le esercita secondo precisi canoni ermeneutici (c.d. *“interpretazione giudiziale”*), non delegabile ad altri, né – meno che mai – ad una macchina algoritmica. Si aggiunge, per completezza, che altra attività sicuramente non delegabile ai sistemi di AI è la valutazione dei fatti e delle prove raccolte nel giudizio e, ovviamente, l'adozione del provvedimento finale, quale la sentenza che decide in rito e/o nel merito.

In definitiva si ribadisce il principio che l'*artificial intelligence*, con l'applicazione del machine learning e le nuove tecnologie, può essere molto utile, nel settore giuridico, per velocizzare le procedure e contribuire, in modo determinante, a ridurre i tempi di svolgimento dei processi, eliminando in tal modo anche l'arretrato formatosi per ragioni varie in qualche sede giudiziaria. È questa la c.d. "giustizia predittiva", inteso il termine quale utilizzo di strumenti basati sull'intelligenza artificiale idonea a supportare la funzione giudiziale, analizzando in tempi brevissimi una grande quantità di informazioni, ed in tal modo "prevedere", anche attraverso lo studio del c.d. "precedente", il possibile esito di un giudizio.

L'intelligenza artificiale, la giustizia predittiva, a cui di recente si è aggiunta la "giurimetria", sono fenomeni già presenti in Europa, che il presente d.d.l. (n. 1066 AS) ha inteso affrontare per la realtà italiana, ponendo precise regole per determinati settori, tra cui quello giuridico. Si chiarisce poi che la giurimetria, nel quadro della calcolabilità del diritto, rappresenta l'applicazione dell'informatica nel diritto, e si occupa in particolare della misurazione e delle analisi dei fenomeni giuridici attraverso metodi quantitativi. Essa viene utilizzata per esaminare le sentenze già pubblicate, evidenziando modelli e tendenze, per fornire indicazioni sulla possibilità di esiti analoghi in fattispecie simili.

Attualmente la giustizia predittiva, la giurimetria e la *cyber* giustizia, intesa quest'ultima quale applicazione di sistemi automatizzati (*hardware* e *software*) e di intelligenza artificiale, con la presenza di robot ed algoritmi decisionali nel campo della giustizia, sono molto diffusi in USA e negli altri Stati del *common law*, il cui modello di Ordinamento giuridico – com'è noto – è basato in prevalenza sui "precedenti giurisprudenziali". Ciò a differenza dei sistemi di *civil law*, a cui appartiene l'Ordinamento italiano, che si basano invece sulla codificazione e su leggi ed atti normativi. Infatti nell'Ordinamento italiano la rilevanza del "precedente giudiziario" ha un valore ben preciso e limitato, anche perché nei sistemi di *civil law* la giurisprudenza non rientra tra le fonti del diritto e il giudice si basa, nel decidere una causa, principalmente sulla normativa codicistica, essendo egli soggetto solo alla legge (art. 101 Cost.) non vincolato da pronunce passate su casi analoghi, ad eccezione che la pronuncia provenga da Organi autorevoli, come le Sezioni unite della Corte di cassazione, *ex art.* 306-bis c.p.c. o anche *ex art.* 384 c.p.c., in cui la Corte Suprema in funzione nomofilattica enuncia il "principio di diritto" da seguire e sul quale le altre Sezioni devono uniformarsi.

Si stabilisce (art. 15 d.d.l.), infine, che le cause che hanno ad oggetto il funzionamento di un sistema di intelligenza artificiale, sono attribuite – nell'ambito della giurisdizione del giudice ordinario – alla "competenza esclusiva" del Tribunale civile, modificando in tal senso l'art. 9, c. 2, c.p.c.

- Cybersicurezza nazionale. Il d.d.l. (n. 1066 AS) in esame stabilisce che all'Acn (acronimo di "Agenzia per la cybersecurity nazionale), Agenzia di nuova istituzione, viene affidato il compito di garantire l'attuazione della normativa nazionale e dell'Unione europea in materia di IA. Ciò per assicurare che tale attuazione avvenga senza pericoli o insidie, per uno sviluppo sicuro in Italia, anche a mezzo di partenariato pubblico-privato, delle tecniche di intelligenza artificiale, intesa quest'ultima quale risorsa di importanza nazionale.

Detta Agenzia, istituita con d.l. n. 82 del 14 giugno 2021, persegue il fine di rafforzare gli aspetti di sicurezza e resilienza cibernetiche nello spazio informatico. In particolare previene e combatte gli attacchi cibernetici da parte di *hacker* verso obiettivi sensibili, pubblici e/o privati, di rilevanza strategica nazionale (esempio ministeri, aziende strategiche, banche, ecc.), combattendo azioni di contrasto e l'utilizzo di *virus*, quali "*social engineering*", applicazioni *web exploit*, attacchi alla "catena di approvvigionamento *cyber*", che negli ultimi tempi sono aumentati in modo esponenziale, di solito colpendo la *supply chain* del *software*, oppure attaccando con *virus* quali i "*Distributed Denial of Service*", oppure i "*mon in the middle*".

L'Acn svolge anche un'ulteriore funzione molto importante: tutela in via preventiva la sicurezza informatica, attraverso l'informazione divulgativa in merito ai pericoli di attacchi *hacker*, sollecitando i soggetti interessati a mantenere un alto il livello di attenzione sulla protezione delle infrastrutture informatiche.

4. La Strategia nazionale; le due nuove Autorità istituite e le azioni di promozione programmate.

Il d.d.l. (n. 1066 AS) ha anche previsto (cfr. artt. 17, 18 d.d.l.), sul piano programmatico, una "strategia nazionale" che garantisca una collaborazione tra pubblico e privato, coordinando a tale riguardo le iniziative della pubblica amministrazione con analoghe attività avviate da investitori privati attratti da incentivi economici, sconti fiscali e bonus per lo sviluppo imprenditoriale ed industriale del settore dell'IA. Il risultato di tale collaborazione è previsto che sia trasmesso alle Camere parlamentari per una valutazione e per l'adozione di misure integrative e correttive al riguardo, al fine di incrementare la ricerca e la realizzazione di risultati positivi.

A rendere più incisivo l'intervento nel settore specifico, sono state istituite due Autorità nazionali per l'IA che – mediante il coordinamento della Presidenza del Consiglio dei ministri – sono preposte alla progettazione ed all'aggiornamento della strategia nazionale per l'IA. Tali nuove Autorità sono costituite dall'Agenzia per l'Italia digitale (AgId) e dall'Agenzia per la cybersicurezza nazionale (Acn), con compiti e funzioni diverse, pur perseguendo il fine comune di garantire l'applicazione e l'attuazione della normativa nazionale dell'Unione europea in materia di *Artificial Intelligence*.

Le due Agenzie suddette, ciascuna nell'ambito della propria competenza, sono preposte alla istituzione ed alla gestione congiunta di spazi di sperimentazione con funzioni di controllo e supervisione, nel settore che realizzano

in Italia sistemi di intelligenza artificiale, conformi alla normativa italiana ed ai principi dell'AI Act approvato di recente dal Parlamento europeo.

Una autonoma valenza, rispetto alle specifiche competenze assegnate alle citate due Agenzie nazionali (Agid e Acn), residua in capo al Garante per la Protezione dei dati personali, di cui però vengono soltanto richiamate le competenze, ma senza attribuzione di compiti specifici nel settore dell'IA.

In particolare l'Agenzia per l'Italia digitale (Agid) è preposta al controllo sulla promozione dell'innovazione e dello sviluppo tecnologico in Italia, curando le procedure di valutazione, accreditamento e monitoraggio delle soluzioni di sistemi di AI destinate al mercato ed assicurandosi che detti sistemi di AI rispettino i criteri di trasparenza, sicurezza ed affidabilità necessari per l'ingresso nel mercato dell'Ue, attraverso ispezioni e con il connesso potere di applicare sanzioni nell'ipotesi di accertata violazione della normativa di settore vigente.

Invece l'Agenzia per la cybersicurezza nazionale (Acn) è preposta, come accennato in precedenza, alla sicurezza nazionale nello spazio cibernetico, contro potenziali minacce di attacchi da parte di *hacker*, italiani e stranieri, verso obiettivi sensibili ed a difesa della sicurezza nazionale e della privacy dei cittadini.

La programmata collaborazione tra Agid e Acn ha il fine ultimo di creare un ambiente controllato e sicuro per lo sviluppo dell'AI, favorendo anche la sperimentazione e la innovazione, in cui i sistemi di intelligenza artificiale, con le innovazioni rivoluzionarie che attua in tutti i settori della vita dei cittadini, si colloca come motore accelerato di crescita economica e di progresso tecnologico, ma a condizione che siano rispettati i principi etici e sociali e di democrazia voluti dalla Costituzione, in sintonia con quelli analoghi dell'Unione europea.

È prevista poi una normativa speciale (art. 20 d.d.l.) per agevolare il c.d. “rientro dei cervelli dall'estero in Italia”, prevedendo per coloro che “hanno svolto un'attività di ricerca anche applicata nell'ambito delle tecnologie dell'intelligenza artificiale” le stesse agevolazioni fiscali ed economiche di cui al d.l. 27 dicembre 2023, n. 209.

Inoltre, nel “piano didattico personalizzato” (Pdp) delle Scuole superiori per studenti e studentesse ad alto potenziale cognitivo sono inserite attività volte alla acquisizione di ulteriori competenze con esperienze di apprendimento presso le istituzioni della formazione superiore. In sintesi, vi è un concreto impegno del Governo a rafforzare i percorsi di formazione alle superiori, in Università e nelle *Its Academy* (acronimo di Istituti tecnici superiori) a sostenere attività di ricerca e di trasferimento tecnologico, ossia a studiare le tecniche algoritmiche dell'IA, riconoscendo agli studenti meritevoli “crediti formativi” nell'ambito del citato “piano didattico personalizzato” (Pdp). Ciò per superare il gap, rispetto ad altri Stati (USA, Cina, Canada, India, Russia, Giappone), per carenza di tecnici ad alto potenziale cognitivo nel settore dell'IA, della cybersecurity, del calcolo quantistico, delle telecomunicazioni e delle tecnologie ad esse pertinenti.

Inoltre, sono previsti investimenti nei settori di IA, della cybersecurity, del calcolo quantistico e del *quantum computing* per l'importo di un miliardo di euro, onde favorire le telecomunicazioni (art. 2 d.d.l.), lo sviluppo, la crescita ed il consolidamento delle imprese operanti in tali settori. Tali investimenti possono realizzarsi per il ruolo importante assegnato a “Cdp Venture Capital sgr”, s.p.a. – Fondo nazionale innovazioni (Fni), di cui è presidente Francesca Bria. A detta società, costituita dal Ministero delle imprese e del made in Italy (MimIt), viene affidata la gestione di un miliardo di euro da dedicare ad investimenti di venture capital. Tale somma verrà utilizzata in questo settore attraverso *start-up* e Pmi innovative, con sede in Italia, operanti in settori chiave, quali AI, cybersecurity, *quantum computing*, telecomunicazioni, 5G, mobile Edge computing e web3. I fondi assegnati possono essere utilizzati sia per la fase di seed (sperimentazione), che per quella di *scaleup* (sviluppo del prodotto).

Infine, per le imprese di dimensioni più grandi, diverse dalle Pmi (acronimo di piccole e medie imprese), gli investimenti previsti possono essere utilizzati per creare i c.d. “campioni nazionali”, ed anche per finanziare “la creazione di poli di trasferimento tecnologico e programmi di accelerazione”. Si riconosce in tal modo l'importanza ed il potenziale trasformativo di queste tecnologie per l'economia e le imprese italiane.

Si precisa infine che detti investimenti possono effettuarsi anche mediante la istituzione di uno o più fondi appositamente dedicati, oppure attraverso coinvestimenti di altri fondi gestiti dal Ministero delle imprese e del made in Italy (MimIt) attraverso la propria società Cdp Venture Capital sgr, s.p.a.

5. La tutela per gli utenti italiani e le modifiche normative in materia di diritto d'autore.

Il provvedimento governativo ha poi apportato modifiche (cfr. art. 23 d.d.l.) al d.lgs. 8 novembre 2021, n. 208, con la previsione – nell'ambito dei servizi di media audiovisivi – di misure per favorire la identificazione ed il riconoscimento dei sistemi di IA nella “creazione di contenuti testuali, fotografici audiovisivi e radiofonici”. Ciò al fine di evitare che siano presentati “come reali dati, fatti o informazioni che invece non sono tali”, in quanto sono stati creati tramite sistemi di IA. Si richiede che tali eventi siano indicati con un “contrassegno” con l'acronimo “IA”, ossia un apposito segno di riconoscimento, onde contrastare efficacemente i c.d. *deepfakes*.

Tale contrassegno deve essere ben visibile e riconoscibile dagli utenti, sia all'inizio della trasmissione, sia alla fine della medesima trasmissione e/o programma, ed anche ad ogni ripresa di esso a seguito di interruzione pubblicitaria. Si tratta di un bollino, che identifichi i contenuti realizzati o modificati dall'IA, da inserire nei canali TV, radio e piattaforme di *streaming*. Si aggiunge che la legge sul *copyright* tutela i prodotti degli algoritmi “nel

caso in cui il contributo umano nell'ideazione e realizzazione dell'opera tramite l'uso dell'algoritmo sia creativo, rilevante e dimostrabile", punendo gli abusi in tal senso con una pena sanzionatoria aumentata.

Inoltre, sono previste (art. 24 d.d.l.) anche significative modifiche anche alla l. 22 aprile 1941 n. 633, a tutela del diritto d'autore delle opere generate dall'intelligenza artificiale. In tal modo si risolve una questione molto discussa, in Italia ed all'estero, in merito alla definizione di "bene giuridico protetto". Il d.d.l. in esame include in tale definizione anche l'opera dell'impegno umano, che sia stato creato "con l'ausilio dell'intelligenza artificiale, purché il contributo umano sia creativo, rilevante e dimostrabile". È onere a carico dell'azienda, che utilizza sistemi di IA per la creazione di opere, rendere "dimostrabile" il contributo dell'uomo sotto il profilo qualitativo, precisandosi che adempiere a tale requisito a volte non è di facile soluzione sotto il profilo probatorio.

È programmato infine che la riproduzione e l'estrazione di opere o altri materiali attraverso modelli e/o sistemi di IA, anche generativa (esempio Chat GPT), sono subordinate alle disposizioni di cui agli artt. 70-ter e 70-quater della citata l. n. 633/1941, al fine di assicurare l'identificazione delle opere e degli altri materiali il cui utilizzo non sia espressamente riservato ai titolari del diritto d'autore. Attualmente è previsto un sistema di *opt-out* che di fatto assegna ai titolari dei diritti sui contenuti *on-line* l'onere di escluderne l'impiego nei "data set" utilizzati per l'addestramento.

6. La tutela penale. Un nuovo reato: l'art. 612-quater c.p.

Il d.d.l. (n. 1066 AS) ha infine modificato (cfr. art. 25 d.d.l.), al fine di predisporre una adeguata tutela penale, nel caso che la normativa suindicata sia violata, l'art. 61, c. 1, del vigente c.p., prevedendo un aumento di pena per i reati connessi mediante l'impiego di intelligenza artificiale, quando sia dimostrato che gli stessi, per la loro natura o per le modalità di utilizzo, abbiano costituito "*mezzo insidioso*", oppure quando il loro impiego "*abbia ostacolato la pubblica o la privata difesa o aggravato le conseguenze del reato*".

Inoltre, è stato previsto un ulteriore aggravamento di pena per il soggetto che, mediante la diffusione di prodotti di IA, abbia tentato di alterare i risultati delle competizioni elettorali, introducendo nel vigente c.p. un nuovo reato, previsto e punito dall'art. 612-quater c.p., per il quale è prevista la pena della reclusione da sei mesi a tre anni, con l'aggravamento della reclusione da uno a cinque anni "se dal fatto deriva un danno ingiusto". Pur dovendo attendere il testo definitivo, da una prima analisi della fattispecie emerge che si tratta di un "reato comune", nel senso che può essere commesso da chiunque, collocato tra i delitti contro "la libertà individuale", e più specificamente "libertà morale", che rappresenta il bene giuridico tutelato. Quanto alla condotta, essa integra una fattispecie che comporta comunque "la falsificazione o l'alterazione mediante impiego di IA". Si tratta di un reato punibile, per l'elemento soggettivo, a titolo di dolo (generico), che consiste nella volontà manifestata di arrecare un danno ad una persona, ingannandola, con artifici vari. Il delitto è punibile "a querela della persona offesa". Tuttavia, si procede d'ufficio qualora il fatto sia connesso con altro delitto, per il quale si deve procedere d'ufficio, oppure se è commesso nei confronti di persona incapace, per età o per infermità, oppure nei confronti di una pubblica autorità a causa delle funzioni esercitate.

Oltre alla creazione di tale nuovo delitto, vengono modificati, dal punto di vista sostanziale, alcuni articoli del codice penale (artt. 501, 612-ter, 640-bis e 640-ter), l'art. 2637 del c.c., e l'art. 171 della legge sul diritto d'autore.

Infine, il Governo si riserva, tra l'altro, di designare, in base all'AI Act di recente approvato dal Parlamento dell'Unione europea, un'Autorità di vigilanza sul mercato dell'intelligenza artificiale, riservandosi di emettere, entro il termine di un anno, "uno o più decreti legislativi per definire organicamente la disciplina nei casi di sistemi di intelligenza artificiale per finalità illecite".

7. Il G 7 svolto dal 13 al 15 giugno 2024 in Italia a Borgo Egnazia di Fasano (BR), con l'intervento di Papa Bergoglio, che ha indicato il ruolo etico dell'Artificial Intelligence.

L'esame del d.d.l. (n. 1066 AS) va integrato con un doveroso riferimento alla riunione del G.7, che si è svolta a Borgo Egnazia, località in Comune di Fasano (BR), dal 13 al 15 giugno 2024, nella splendida cornice dell'omonimo Resort di lusso sul Mare Adriatico, il cui focus è stato proprio l'intelligenza artificiale.

La riunione del G7, a presidenza italiana (3), con la partecipazione – oltre all'Italia, USA, Canada, Francia, Regno Unito, Germania e Giappone – ha compreso anche il Pontefice Francesco Bergoglio, per la prima volta ad un G7, l'Unione Europea, con la Presidente Ursula von der Leyen, il Presidente dell'Ucraina, Volodymyr

(3) F. Del Vecchio, A.I. G7. Papa Francesco ha parlato di IA. Ecco cosa ha detto (14 giugno 2024), in <www.wired.it>; M. Muolo, Il Papa al G7: l'IA resti al servizio dell'uomo; serve l'etica (14 giugno 2024), in <www.avvenire.it>; R. Marino, Papa Francesco al G7: IA. affascinante, ma tremenda (14 giugno 2024), in <www.quotidiano.net>; Redazione Vatican News, Papa Francesco parteciperà alla sessione del G7 sull'intelligenza artificiale (26 aprile 2024) in <www.vaticannews.va>; Redazione Ansa, Meloni, il Papa alla sessione del G7 sulla intelligenza artificiale (27 aprile 2024), in <www.ansa.it>; Redazione La Sicilia, G7 in Puglia, Virgilio influencer con l'intelligenza artificiale (20 maggio 2024), in <www.lasicilia.it>; Mediazione ESG360, Rome Call for A.I. Ethics 2023: etica e innovazione devono superare anche i confini religiosi (10 gennaio 2023), in <www.esg360.it>; A. Mazzuca, G7 Lavoro: I temi ed il programma di lavoro. Al centro l'Intelligenza artificiale e la sicurezza del lavoro (16 febbraio 2024), in <www.insic.it>

Zelensky, e la Direttrice Generale del Fondo Monetario Internazionale, Kristalina Georgieva, nella sessione aperta ai Paesi non membri, la c.d. “Ontreach”.

Com'è noto, il G7, nato nella prima metà degli anni '70 a seguito della crisi del sistema di Bretton Woods e della crisi energetica del 1973, come foro informale di coordinamento in materia economica e finanziaria, riunisce periodicamente i sette Paesi altamente industrializzati, a cui si è aggiunta di recente, per la sua crescente importanza politica ed economica, anche l'Unione europea. I Paesi che fanno parte del G7 sono uniti da una serie di principi e valori comuni, quali libertà e democrazia, *rule of law*, il rispetto dei diritti fondamentali dell'uomo.

La sessione del G7 del 13-15 giugno 2024 svoltasi a Borgo Egnazia è iniziata in una atmosfera di grande fascino e novità, in quanto la società Carraro Lab, specializzata nello sviluppo di piattaforme e tecnologie nei settori più avanzati dei media digitali – con l'IA, attraverso la metodologia della “*re-generative art*” – ha ricostruito la figura del poeta Virgilio, rievocando il suo viaggio ad Egnazia, lungo la via Appia, nel 37 a.c..

Il poeta, come un influencer digitale, ha ricevuto gli ospiti presenti per la sessione del G7 all'insegna dell'ospitalità, che affondava, con tale rievocazione storico culturale, le sue radici nel mondo antico. Il volto del poeta Virgilio è stato ricostruito in 3D, sulla base delle sculture esistenti, scegliendo quella ritenuta verosimile, con la rievocazione del suo viaggio ad Egnazia nel 37 a.c., e la creazione in tal modo di un Case History applicativo dell'intelligenza artificiale, quale focus della sessione del G7 2024.

Infatti, tra i temi trattati nella sessione del G7, a presidenza italiana – a parte i vari argomenti di natura politica, che avevano una loro importanza strategica per gli equilibri geo-politici dei Paesi partecipanti e che sono stati esaminati anche con incontri bilaterali tra i grandi della terra – era riservato uno spazio importante all'*Artificial Intelligence*, il cui focus erano le enormi opportunità, ma anche i tanti rischi ed i pericoli ad essa connessi, se non regolamentata da principi etici condivisi aventi natura cogente. Quindi, in questa sessione del G7 ha trovato un suo autonomo spazio il tema della Intelligenza Artificiale, quale nuova realtà tecnologica, straordinaria e complessa, dalle applicazioni innumerevoli, ma anche con rischi e pericoli concreti per i diritti fondamentali dell'uomo e per i valori etici di una società moderna e democratica.

In tale occasione, Papa Francesco ha esposto il pensiero centrale contenuto nell'importante documento costituito dal “*Rome call for Artificial Intelligence Ethics*”, elaborato fin dall'anno 2020. Il Pontefice, alla presenza dei grandi della terra, ha ampiamente valorizzato il progetto promosso dalla Santa Sede sulla intelligenza artificiale con la citata “*Rome Call*”, dando un contributo decisivo alla definizione di un quadro regolatorio, etico e culturale, sull'utilizzo dell'intelligenza artificiale nel mondo. Si ricorda che la Santa Sede dal 2020 aveva avviato un percorso culturale per la creazione della c.d. “*Algoetica*”, ossia una riflessione etica sull'uso degli algoritmi. Tale percorso, iniziato con la firma della “*Rome Call for AI Ethics*” in seno alla Pontificia Accademia per la vita, si è ampliato con la collaborazione di altri enti internazionali e società, quali la Fao, Ibm, Microsoft ed il Ministero dell'innovazione italiano, perseguendo il fine di un approccio etico allo sviluppo dell'*Artificial Intelligence*. Tale progetto è stato poi esteso e recepito anche dalle fedi musulmana ed ebraica, rendendo in tal modo universale l'obiettivo multiculturale e globale di garantire un approccio etico all'intelligenza artificiale, senza trascurare l'obiettivo della giustizia e della pace nel mondo.

La “*Rome Call for AI Ethics*” rappresenta una cornice valoriale di rilevante importanza per l'applicazione etica e sociale di questa nuova tecnologia, indicando come essenziali nell'utilizzo dell'IA i principi fondamentali di trasparenza, inclusione, responsabilità, imparzialità, affidabilità, sicurezza e privacy. La “*Rome Call*” ha poi ricevuto il consenso anche del “Forum per la Pace” svoltosi di recente a Abu Dhabi negli Emirati Arabi e della Fondazione Vaticana “*RenAlssance*”, istituita nel 2021, con sede nella Città del Vaticano, quale organo etico dell'Accademia della Vita. Si aggiunge che di recente ha aderito alla “*Rome Call for AI Ethics*” anche la multinazionale USA della CISCO System Inc., con sede in California, azienda leader mondiale per l'adozione e l'implementazione dei sistemi di AI, la quale ha sempre agito e pensato in una prospettiva *top-down*, in una *governance* proattiva e nella mitigazione dei rischi, nonché in una prospettiva sull'impatto trasformativo dell'AI nei settori industriali e nella società civile.

L'intervento del Pontefice ha chiarito che l'AI è una grande opportunità per il progresso dell'umanità, ma può costituire un rischio per la democrazia, per l'etica e per i diritti fondamentali dell'uomo, se non sia regolamentata secondo principi e regole giuste ed equilibrate. Il Santo Padre, alla riunione del G7 del 15 giugno 2024 svoltasi a Borgo Egnazia – proseguendo le iniziative assunte al *summit di Hiroshima* nel 2023, quando era stato introdotto un codice etico, avente però carattere volontario, per le aziende e le amministrazioni pubbliche che utilizzano l'AI – ha indicato la strada giusta da seguire, nel senso che è necessario stabilire regole etiche di natura vincolante nei confronti di una tecnologia che, se lasciata senza controllo, può alimentare la disinformazione e porre in serio pericolo la democrazia, la privacy e l'identità digitale dell'uomo.

Papa Francesco ha espresso la necessità che l'*Artificial Intelligence* sia collegata ad una dimensione etica, affermando testualmente: “*il vigoroso avanzamento tecnologico rende l'intelligenza artificiale uno strumento affascinante e tremendo al tempo stesso, ed impone una riflessione all'altezza della situazione*”, sottolineando che: “*parlare di tecnologia è parlare di cosa significhi essere umani e quindi di quella nostra unica condizione tra libertà e responsabilità cioè vuol dire parlare di etica*”. “*Non possiamo nascondere il rischio concreto, poiché insito nel suo meccanismo fondamentale, che l'intelligenza artificiale limiti la visione del mondo ad una realtà*”.

esprimibile in numeri e racchiusa in categorie preconfezionate, estromettendo l'apporto di altre forme di verità e imponendo modelli antropologici, socioeconomici e culturali uniformi. Il paradigma tecnologico incarnato dall'intelligenza artificiale rischia allora di fare spazio a un paradigma ben più pericoloso". Il Papa ha poi evidenziato che: "Di fronte ai prodigi delle macchine, che sembrano saper scegliere in maniera indipendente, dobbiamo aver ben chiaro che all'essere umano deve sempre rimanere la decisione".

Anche l'impiego a fini militari è stato un tema affrontato con molta chiarezza dal Pontefice, il quale ha stigmatizzato che: *"È urgente ripensare lo sviluppo e l'utilizzo di dispositivi, come le c.d. "armi letali autonome", per bandirne l'uso, cominciando già da un impegno fattivo e concreto per introdurre un sempre maggiore e significativo controllo umano. Nessuna macchina dovrebbe mai scegliere se togliere la vita ad un essere umano".*

Papa Francesco – continuando la sua opera profondamente incisiva, iniziata a Lampedusa con i migranti e continuata con l'enciclica *"Laudato Si"* con il tema ambientale – ha ora richiamato l'attenzione dei potenti della terra sull'*Artificial Intelligence*, definita *"strumento affascinante e tremendo al tempo stesso"*, da utilizzare secondo criteri valoriali precisi, ispirati ai principi fondamentali di trasparenza, inclusione, responsabilità, imparzialità, affidabilità, sicurezza e privacy, e comunque sempre nell'intesa *"che all'essere umano deve sempre rimanere la decisione"*.

8. Conclusioni.

Il presente lavoro ha analizzato il d.d.l. (n. 1066 A.S.), di iniziativa governativa, presentato il 23 aprile 2024, con il quale si è inteso disciplinare l'utilizzo dell'intelligenza artificiale, tracciando il quadro regolatorio da applicare negli spazi specifici del diritto interno italiano, in una visione proattiva e cautelativa dell'IA, senza sovrapporsi al Regolamento *"IA Act"*, approvato dal Parlamento Europeo il 23 aprile 2024, ma anzi in perfetta sintonia con i principi da esso sanciti.

L'iniziativa del Governo italiano colloca l'Italia in una posizione primaria nel dibattito globale, nel fissare le regole con le quali l'IA può essere utilizzata in Europa e nel resto del mondo. Il punto focale del d.d.l. (n. 1066 A.S.) è che l'utilizzo di questa straordinaria innovazione algoritmica e tecnologica abbia *"una dimensione antropocentrica"*, nel senso che mantenga l'uomo al centro del processo decisionale, garantendo che la tecnologia e gli algoritmi siano solo un supporto ad un ausilio all'attività dell'uomo, senza sostituirsi ad esso, in una dimensione etica e consapevole, e che il suo utilizzo avvenga sempre secondo principi di correttezza, trasparenza e responsabilità.

Tali propositi, normativamente concretizzati con il d.d.l. (n. 1066 A.S.), hanno trovato una importante conferma nella sessione del G7 tenuto dal 13 al 15 giugno 2024 a Borgo Egnazia (BR), a presidenza italiana, con la partecipazione di Papa Bergoglio, con un *focus* specifico per l'*Artificial Intelligence*, *"strumento affascinante e tremendo al tempo stesso"*, da regolamentare secondo principi di trasparenza, inclusione e responsabilità, in una corretta declinazione dei valori costituzionali, del diritto alla riservatezza e alla privacy.

Un ulteriore e più specifico approfondimento è stato poi svolto – per i profili correlati al diritto pubblico – al ruolo che l'IA è chiamata a svolgere per una migliore funzione della giustizia. In tale importante settore l'utilizzo dei sistemi di intelligenza artificiale si ritiene sia ammessa limitatamente alla *"giustizia predittiva"* (con esclusione della c.d. *"giustizia generativa"*), ossia solo per finalità *"strumentali e di supporto"*, per agevolare e semplificare il lavoro giudiziario con la ricerca di precedenti giurisprudenziali o di articoli di dottrina, favorendo in tal modo lo studio e la soluzione di casi analoghi, quando il giudice deve risolvere una controversia simile ad altre già decise.

L'applicazione del *machine learning* e delle nuove tecnologie algoritmiche può dare il proprio apporto tecnologico per velocizzare le procedure e contribuire a ridurre i tempi di svolgimento dei processi. Nell'ambito del settore giustizia gli algoritmi di IA possono analizzare grandi quantità dei dati giuridici, riguardanti i testi di leggi, precedenti giudiziari e decisioni delle Corti, onde fornire un supporto importante per redigere le decisioni. Quindi l'IA può riprodurre gli esiti di una vertenza, in sede civile, penale, tributaria, contabile sulla base di un numero altissimo di precedenti e tenendo presenti tutti i testi di leggi vigenti, in tempi brevissimi, allo scopo di *"prevedere"*, in modo quanto più possibile certo e preciso, il possibile esito di un procedimento.

È questa la c.d. *"giustizia predittiva"* (4), utilizzabile anche dalla Corte dei conti, nello svolgimento delle delicate funzioni giurisdizionali, di controllo e di consulenza contabile, in sede centrale e periferica.

(4) C. Sorrentino, *Giurimetria e Giustizia predittiva: intelligenza artificiale e Legal Tech* (26 luglio 2023), in <www.diritto.it>; C. Catelli e D. Piana, *Giusto processo e intelligenza artificiale*, Ed. Maggioli, 2019; L. Viola, *Giustizia predittiva: è preferibile un modello deduttivo* (10 marzo 2020), in <www.altalex.com>; F.G. Capitani, *Chi vince e chi perde con l'intelligenza artificiale nel diritto* (4 aprile 2023), in <www.dirittoegiustizia.it>; M. A. Catarozzo, *Giustizia predittiva e Intelligenza Artificiale: quel bisogno antico di prevedere il futuro* (20 maggio 2023), in <www.dirittoegiustizia.it>; G. Ludovici, *Giustizia predittiva deduttiva, funzione giurisdizionale e certezza del diritto* (22 giugno 2022), in <https://officeadvice.it>; A. Traversi, *Giustizia predittiva: quale futuro?* (15 marzo 2023), in <www.altalex.com>.

Il problema dell'utilizzo dell'IA nel settore giustizia è stato affrontato, com'è noto, anche nell'Unione europea con la Cepej (acronimo di *Commission European for the Efficiency of Justice*) (5), istituita nel 2002, la quale ha studiato l'impatto delle tecnologie dell'informazione sui sistemi giudiziari europei. Essa ha svolto, nel 2016, uno studio analitico sul loro uso nei Tribunali europei. Inoltre, nel 2019, è stata redatta la “*Ethics guidelines for trustworthy AI European Commission*”, che è la “Carta etica europea sull'uso dell'IA nei sistemi giudiziari”, con cui è stata posta una specifica attenzione sulle tecniche di analisi automatizzate usate nelle decisioni giudiziarie. La Carta è un documento che, sebbene non vincolante, enuncia principi sostanziali e di metodo validi per i soggetti privati e per le pubbliche autorità. Essa è molto importante in quanto tutela i diritti fondamentali dell'uomo, il principio di non discriminazione e di trasparenza delle tecniche usate per le decisioni giudiziarie e fissa il principio che le parti di un giudizio devono essere sempre informate e poter controllare le proprie scelte rispetto alle decisioni dell'Autorità giudiziaria.

In Italia la c.d. “giustizia predittiva” ha suscitando un interesse sempre maggiore in campo giuridico, ed è utilizzata in modo sempre più diffuso (6), con il fine di prevedere l'esito di un giudizio in ambito civile, commerciale, penale, tributario. Anche la Magistratura contabile utilizza strumenti informatici di ultima generazione per la organizzazione dei propri uffici e gestisce con professionalità le problematiche (7) dell'IA, in relazione alla graduale trasformazione digitale delle strutture dei propri uffici giudiziari, tenuto conto che la digitalizzazione è prevista peraltro dall'art. 6 del d.lgs. n. 174/2016 come novellato dal d.lgs. n. 114/2019 (8). Inoltre, la Magistratura contabile utilizza il sistema informatico “*Giudice*”, che opera in via automatizzata per tutte le attività svolte dalla Corte dei conti, compresi i dati ed i servizi che riguardano l'attività giurisdizionale. Invece il sistema informatico “*Sireco*”, con l'ausilio delle risorse hardware e software, viene utilizzato dalle Sezioni regioni contabili per l'esame dei conti giudiziali.

Quindi può affermarsi che la “giustizia predittiva” svolge, nell'ambito della Magistratura contabile, un ruolo molto importante anche con riguardo alla nomofilachia (9), la cui funzione – com'è noto – è rivolta a garantire la esatta osservanza e l'uniforme interpretazione della legge. Il principio della certezza del diritto, con riguardo alla giurisprudenza del giudice civile, penale, amministrativo e contabile, dopo la riforma della Costituzione del 1999 (l. cost. 23 novembre 1999, n. 2), che ha sancito il principio del “giusto processo” e della sua “ragionevole durata”, ha cambiato anche l'effetto finale della pronuncia di nomofilachia, nel senso che ora la uniformità delle decisioni

(5) C. Barbaro, *Uso dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione dei principi etici condivisi a livello europeo?*, in <www.questionegiustizia.it>; F.C. Castaldo, *Lo Statuto della giustizia digitale nella Carta etica della Cepej* (2 aprile 2021 e aggiornato al 6 giugno 2022), in <www.insiniture.it>; *Decimo rapporto di valutazione sui sistemi giudiziari della Cepej* (9 febbraio 2023), in <www.processopenalegiustizia.it>; C. Barbaro (a cura), *Cepej adottata la Carta etica europea sull'uso dell'intelligenza artificiale (IA) nei sistemi giudiziari* (7 dicembre 2018), in <www.questionegiustizia.it>; M. Castellaneta, *Cepej: piano d'azione sulla digitalizzazione per una giustizia migliore* (16 dicembre 2021), in <www.marinacastellaneta.it/blog/wp>. Con tale articolo si evidenzia che la Cepej, ha pubblicato un documento che punta a favorire l'ambiente digitalizzato, migliorando la qualità della giustizia che deve essere trasparente, collaborativa, umana. La Cepej non discute più se utilizzare i sistemi di intelligenza artificiale nel trattamento di decisioni e di danni giudiziari, ma delle modalità e dei limiti all'utilizzo per tutelare i diritti fondamentali della persona.

(6) In Italia un primo progetto è stato avviato dalla Corte d'appello e Tribunale di Brescia, in collaborazione con l'Università degli Studi di Brescia. È un sistema di “giustizia predittiva” operativo dal novembre 2021, che interessa al momento solo il “Diritto del lavoro” e il “Diritto delle Imprese”, accessibile sul sito web di detti uffici giudiziari, per fornire “dati di certezza e di prevedibilità e contenere la domanda di contenzioso, disincentivando le cause temerarie”, nonché favorire la “circolarità della giurisprudenza tra il primo ed il secondo grado di giudizio”. Un secondo progetto è stato promosso poi dalla Corte d'appello di Venezia, in collaborazione con l'Università degli Studi Cà Foscari di Venezia, Unioncamere del Veneto e la Soc. Deloitte, avente ad oggetto, al momento, le vertenze per “licenziamenti per giusta causa, e prendendo in esame i precedenti del triennio 2019-2021, per rendere conoscibili le sentenze e scoraggiare il contenzioso con scarse prospettive di successo. Un terzo progetto, denominato “Prodigit”, riguarda il settore tributario, avente fini di giustizia predittiva, per prevedere l'orientamento del giudice su questioni tributarie.

(7) In dottrina: V. Tenore, *I robot in giudizio! Considerazioni sull'utilizzo dell'intelligenza artificiale (IA) da parte del Magistrato: abdicare al ragionamento e alla riflessione quotidiana a favore di un robot togato, ovvero un “cretino digitale”, è un bene per la giustizia e per i suoi attori e destinatari?*, in questa Rivista, 2023, fasc. 4; Convegno organizzato dalla Sezione giurisdizionale Regione Emilia Romagna: *Intelligenza artificiale: giustizia predittiva e giudice sono compatibili?*, in data 24 novembre 2023 a Bologna; A. Giordano, *Intelligenza Artificiale, giusto processo e giudizio contabile*, in questa Rivista, 2022, fasc. 3, 27.

(8) Decreto del Presidente della Corte dei conti, 24 maggio 2022 n. 126, relativo a “*ulteriori regole tecniche e operative per lo svolgimento dei giudizi dinanzi alla Corte dei conti mediante le tecnologie dell'informazione e della comunicazione*” (G.U. n. 123 del 27 maggio 2022), quale attuazione delle direttive di cui all'art. 6 c.g.c.

(9) G. Arioli, *Nomofilachia, giustizia predittiva e intelligenza artificiale* (3 novembre 2023), in <www.giustiziainsieme.it>; E. Vincenti, *Massimazione e conoscenza della giurisprudenza nell'era digitale*, in <www.questionegiustizia.it>; M.F. Carriero, *Crisi della legalità e intelligenza artificiale: uno sguardo (comparato) d'insieme*, in *Archivio Penale*, 2024, fasc. 2, 1-36; M.C. Razzano, *La nomofilachia nel codice di giustizia contabile*, in *Dirittoeconti*, 2022, fasc. 1; C. Perago, *Banca dati di merito tra automatismo decisionale e umanità del diritto* (15 maggio 2024) in <www.judicium.it>; R. Blaiotta, *Giustizia, errore, intelligenza artificiale* (19 ottobre 2023), in <www.sistemapenale.it>.

costituisce anche un valido strumento per velocizzare i tempi della giustizia e realizzare il duplice risultato di avere una giustizia, che sia rapida e nel contempo anche giusta.

La funzione di interpretazione uniforme delle pronunce dei giudici contabili, com'è noto, trova la sua disciplina negli artt. 11 e 114 del vigente c.g.c. In particolare, l'art. 11 conferisce alla Sezioni riunite la funzione di enucleare dalle pronunce dei giudici il "principio di diritto", ossia la *regula iuris* di soluzione del caso controverso, concorrendo in tal modo alla generalizzazione dell'interpretazione ed alla uniforme applicazione delle norme di contabilità pubblica.

In questo senso il valore e la forza del "precedente giurisprudenziale" assume rilevanza sul piano della persuasività, quale aspirazione ad una tendenziale stabilità, in armonia con la funzione nomofilattica. Quindi, anche nella giustizia contabile, detta funzione si esprime nell'ottica valoriale della certezza giuridica, costituendo "il precedente" il criterio e la misura della prevedibilità e calcolabilità rispetto alla decisione di casi futuri.

Importante al riguardo è l'Ufficio del Massimario della Corte dei conti, che provvede, tra l'altro, all'esame delle sentenze pronunciate dalle Sezioni riunite, per estrarne in forma di "massime" i principi giuridici enunciati, che riguardano i precedenti giurisprudenziali. Detto Ufficio cura anche la massimazione di sentenze e deliberazioni delle Sezioni di controllo emesse dalla Corte, provvedendo all'aggiornamento e gestione dell'archivio della giurisprudenza contabile sulla base della banca dati del sistema informativo della Corte e di eventuali altre banche dati a questa collegate.

Alla luce delle considerazioni che precedono, può affermarsi quindi – in via conclusiva – che l'intelligenza artificiale, "strumento affascinante e tremendo al tempo stesso", secondo la pertinente definizione di Papa Bergoglio, risulta ben disciplinata nel d.d.l. (n. 1066 A.S.), essendone previsto un "*utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica*", in linea con i principi già approvati dall'Unione europea. In tal modo vengono tracciate le regole per la ricerca, lo sviluppo e la sua corretta applicazione in Italia, peraltro migliorabile, con il contributo dialettico del Parlamento, in sede di approvazione del testo definito (10).

Si ritiene che l'utilizzo dell'intelligenza artificiale nei sistemi giudiziari, compreso quello dinanzi alla Corte dei conti, debba necessariamente rispettare *in primis* i principi fondamentali, riferibili direttamente a situazioni giuridiche soggettive, sostanziali e processuali, connesse alla tutela giudiziaria a garanzia del giusto processo, del diritto di difesa e della imparzialità del giudice. Viene, poi, tutelata anche la trasparenza, la qualità e la sicurezza del cyber spazio in cui sono allocati i "casi" giudiziari da catalogare e schematizzare nelle banche dati predisposte alla loro soluzione. Ciò per un sostanziale rispetto del principio di non discriminazione, della qualità e sicurezza dei dati, di trasparenza, imparzialità ed equità e di pieno controllo dell'operatore. Sulla base di tali garantistici presupposti la "giustizia predittiva" può essere utilizzata nel settore giustizia, a supporto all'attività del giudice, unico soggetto a decidere la giusta sentenza.

La c.d. "umanità" della decisione è legata sicuramente alla valutazione degli atti del processo, delle prove raccolte, ma anche di altri elementi, tra cui l'intuizione del giudice, la sua capacità di cogliere la diversità del "caso concreto", i singoli aspetti sociali e culturali in cui vi è sorta e maturata la vicenda per cui è causa, riuscendo a rendere – anche sulla base di tali ulteriori elementi – una "giusta giustizia". È questa la cornice valoriale necessaria per l'utilizzo dell'intelligenza artificiale nel settore giustizia, in cui l'uomo resta al centro del processo decisionale, mentre la tecnologia e gli algoritmi dell'AI costituiscono solo il supporto e l'ausilio all'attività dell'uomo, ma senza sostituirsi ad esso, deputato sempre e comunque ad assumere le decisioni finali, aventi anche connotati di natura etica e sociale. Il tal modo l'IA costituisce un tassello operativo importante per costruire un mondo tecnologico migliore.

* * *

(10) Si rappresenta, per completezza di esposizione, che in data 19 ottobre 2023 è stato presentato il d.d.l. (S n. 917), a firma del Sen. A. Nicita del P.D., recante "*Misure sulla trasparenza dei contenuti generati da intelligenza artificiale*", con il quale si richiede che chiunque "crei, pubblichi o diffonda, in tutto o in parte, prodotti generati da AI, deve farlo con "etichette o filigrane ben distinguibili", con modalità definite dall'Autorità per le garanzie delle comunicazioni (Agcom), previa approvazione di un apposito Regolamento che contenga idonei strumenti di "segnalazioni e rimozione di contenuti", se non conformi a legge. A preoccupare sono i c.d. *deep fake*, ossia i contenuti audio, immagini, o video manipolati o sintetizzati, in modo da farne apparire falsamente autentici o veritieri i contenuti, utilizzando tecniche di AI, di cui fanno parte l'apprendimento automatizzato (*machine learning*) e profondo (*deep machine learning*).

Un secondo d.d.l. (S n. 908), è stato poi presentato sempre al Senato, a firma del Sen. A. Nicita ed altri parlamentari del P.D., il quale prevede l'introduzione di una "legge annuale per il settore digitale, che faccia il punto sullo sviluppo del settore digitale e affronti i nodi irrisolti in detto settore. La proposta prevede anche l'istituzione di un Comitato per lo sviluppo digitale e la regolazione composto da un rappresentante di Agcom (Autorità garante della concorrenza e del mercato), Arera (Autorità di regolazione dei trasporti), il Garante della privacy, l'Agenzia della cybersicurezza nazionale e l'Agenzia per l'Italia digitale. Il compito di tale Comitato è quello di predisporre, entro il 30 settembre di ogni anno, una "relazione" sullo stato dello sviluppo digitale da trasmettere al governo per i conseguenti provvedimenti da assumere.

I due d.d.l. intendono affermare la natura di neutralità tecnologica e politica dell'intelligenza artificiale, distinguendo su piani diversi tecnologia logaritmica ed informazione, quale espressione dell'interazione dell'uomo, onde facilitare l'accesso di imprese, utenti, consumatori e lavoratori ad un "ecosistema digitale che ne tuteli i diritti e che diffonda equamente i suoi benefici".

LA C.D. “INTELLIGENZA ARTIFICIALE”: PERSONALITÀ ELETTRONICA, IMPUTAZIONI E RESPONSABILITÀ TRA DIRITTO CIVILE, PENALE ED AMMINISTRATIVO

di Pasquale Fava (*)

Abstract: L'IA è entrata nella vita dell'uomo ponendo nuove sfide giuridiche, tra cui l'ammissibilità di una "personalità giuridica elettronica" e l'identificazione delle tecniche di imputazione di atti favorevoli e sfavorevoli e delle correlative responsabilità civili, penali e amministrative. Venendo in rilievo un utilizzo globale di nuove tecnologie, l'approccio del giurista all'IA non può che essere multilivello e comparato. Dando conto della più innovativa casistica pretoria ed amministrativa (in materia di veicoli a guida autonoma nonché di proprietà industriale ed intellettuale), lo studio verifica l'impatto dell'IA nel diritto civile, penale ed amministrativo, partendo dal quadro giuridico eurounitario (regolamento "IA Act 2024" e proposte di direttive in materia di responsabilità connessa all'utilizzo dell'IA) e giungendo all'esame del disegno di legge governativo dell'aprile 2024, tenendo conto dei principi affermati dalla giurisprudenza e dal Legislatore (strumentalità, riserva di umanità, tutela, accessibilità e trasparenza).

AI has entered human life by posing new legal challenges including the admissibility of an "electronic legal personality" and the identification of techniques for attribute favorable and unfavorable acts and related civil, criminal and administrative liabilities. As global use of new technologies comes to the fore, the jurist's approach to AI can only be multilevel and comparative. Giving account of the most innovative case law and administrative praxis (in the field of self-driving vehicles as well as industrial and intellectual property rights), the study verifies the impact of AI in civil, criminal and administrative law, starting from the European Union legal framework ("IA Act 2024" Regulation and proposed Directives on liability related to the use of AI) and arriving at the examination of the April 2024 Government Bill, taking into account the principles affirmed by Italian case law and the Legislator (instrumentality, reservation of humanity, jurisdictional protection, accessibility and transparency).

Sommario: 1. *Le “sfide giuridiche” poste dall’IA.* – 2. *Intelligenza artificiale e umana, input e output, strumentalità e riserva di umanità.* – 3. *La c.d. “personalità giuridica elettronica” e il c.d. “Robot-schiavo”: l’IA resta una res.* – 4. *L’imputazione all’IA di atti favorevoli. L’analisi delle controversie in materia di tutela della proprietà intellettuale e industriale. L’impossibilità di attribuire direttamente all’IA la qualifica di inventore e/o autore.* – 5. *IA e profili problematici di responsabilità civile: il caso dei veicoli a guida autonoma. L’applicazione dei criteri di imputazione previsti dal Codice civile e i rapporti con la disciplina eurounitaria approvata e approvanda.* – 6. *Le linee di politica legislativa europea: l’EU IA Act 2024.* – 6.1. (Segue) *L’adeguamento della direttiva sulla responsabilità del produttore alle problematiche poste dall’IA: il rapporto tra disciplina europea sulla responsabilità del produttore, altre normative eurounitarie specifiche e discipline nazionali contrattuali e extracontrattuali.* – 6.2. (Segue) *La proposta di direttiva sulla responsabilità per danni derivanti dall’output dell’IA: l’integrazione dei sistemi nazionali fondati sulla colpa.* – 6.3. (Segue) *Il quadro normativo eurounitario nel complesso: conclusioni.* – 7. *Il disegno di legge governativo italiano dell’aprile 2024: campo di applicazione, definizioni e principi generali. Amministrare e giudicare con algoritmi e IA?* – 7.1. (Segue) *IA e tutela della proprietà intellettuale.* – 7.2. (Segue) *IA e diritto penale: aggravanti e il nuovo reato di illecita diffusione di contenuti generati o manipolati artificialmente.*

1. *Le “sfide giuridiche” poste dall’IA*

L'IA è entrata nelle nostre vite e con essa il giurista è chiamato a confrontarsi onde adeguare categorie giuridiche e colmare i vuoti di regolazione e tutela conseguenti all'utilizzo delle nuove tecnologie. Come avviene sovente con riguardo allo sviluppo di queste ultime è necessario procedere con un approccio multilivello e comparato tenendo conto che in aprile e maggio 2024 il Governo italiano e l'Unione europea hanno approvato rispettivamente il disegno di legge in materia di IA e c.d. IA Act.

Le questioni giuridiche che ha posto, pone e porrà l'intelligenza artificiale tagliano trasversalmente tutte le principali materie (civile, penale, amministrativo e discipline collegate).

L'intelligenza artificiale ha sollevato non solo importanti problematiche giuridiche, ma anche temi di natura etica.

(*) P. Fava è Consigliere della Corte dei conti.

L'approccio di tutte le giurisdizioni, a livello globale, appare, da una parte, aperto allo sviluppo della tecnologia e, quindi, interessato allo sviluppo di questo "nuovo mercato" di prodotti e servizi, ma, allo stesso tempo, risulta essere molto attento alla tutela dei diritti fondamentali con particolare riguardo a quelli della persona umana (profilo che sta molto a cuore alle Istituzioni EU). Ecco perché, ad esempio, il bilanciamento tra questi due interessi contrapposti, ha portato alla recente approvazione di un disegno di legge governativo in materia di intelligenza artificiale, il quale tocca tante tematiche, disciplinando aree lasciate "libere" dal diritto eurounitario *in fieri*: tenendo conto che anche le Istituzioni europee stanno elaborando atti di regolazione dopo uno studio tecnico-giuridico dell'IA, il Legislatore italiano ha fatto la scelta di intervenire, ragionando in una prospettiva *de iure condendo*, in spazi non coperti dall'approvando diritto eurounitario.

Dottrina e giurisprudenza, ma anche Pubbliche amministrazioni (si pensi, ad esempio, agli Uffici di registrazione di brevetti), hanno iniziato ad affrontare queste questioni alla ricerca di soluzioni ragionevoli e condivisibili.

Si tratta anche di questioni di teoria generale del diritto (si pensi al problema della "personalità giuridica elettronica"), ma con rilevanti ricadute di ordine pratico (per esempio con riguardo al tema delle imputazioni giuridiche degli atti e delle responsabilità).

L'approccio che segue il Legislatore italiano è molto prudente, considerando le c.d. IA (intelligenze artificiali) in una prospettiva di strumentalità, con salvezza della c.d. "riserva di umanità" (si pensi, ad esempio, all'utilizzo dell'intelligenza artificiale nell'ambito della pubblica amministrazione o della giustizia). L'intelligenza artificiale è così vista come risorsa o mezzo che può aiutare la pubblica amministrazione e i giudici ad esercitare i propri poteri.

L'intelligenza artificiale apre un orizzonte sconfinato all'interprete.

Dallo studio di precedenti giurisprudenziali, provvedimenti amministrativi e studi dottrinali italiani e stranieri emergono una serie di questioni ricorrenti.

Sicuramente, il primo problema in ordine logico, dal punto di vista teorico, è quello dell'attribuzione di una personalità elettronica alla macchina e all'IA (specie se generativa). Ci si chiede, cioè, se sia possibile attribuire una personalità giuridica ad una macchina, in modo da rendere quest'ultima *soggetto di diritti* e, dunque, centro di imputazione e riferimento di interessi e di rapporti giuridici. La dottrina si è chiesta se sia possibile *de iure condito* rispondere positivamente al quesito oppure, nella negativa, se *de iure condendo* la soluzione sia auspicabile o giuridicamente ammissibile.

Vi è, dunque, il tema della riferibilità degli atti alla macchina e all'IA, problematica che si snoda in relazione sia agli atti favorevoli che a quelli sfavorevoli.

Si pensi, con riguardo ai primi, alla tutela dell'autore di un'opera dell'ingegno, o alla tutela riconosciuta all'inventore di un *quid* che sia brevettabile: ci si chiede se questa tutela sia riferibile direttamente alla macchina o all'IA (specie se generativa).

In relazione agli atti sfavorevoli questa imputabilità potrebbe anche determinare delle conseguenze sfavorevoli, in quanto la macchina potrebbe essere utilizzata come schermo da chi, invece, intenda realizzare condotte illecite (a rilevanza penale, civile o amministrativa): emerge il problema della riferibilità degli atti illeciti a chi abbia progettato il *software*, sia proprietario dell'*hardware* o a chi abbia utilizzato la macchina o l'IA. Altro problema connesso è quello della responsabilità derivante dai danni determinati dall'*output* generato dalla macchina o dall'IA (fattispecie ben nota è quella dei danni determinati da veicoli a guida autonoma).

Ecco i principali problemi sino ad ora posti dall'intelligenza artificiale.

L'operatore del diritto è chiamato a riflettere per inventare soluzioni regolative o pretorie accettabili al fine di risolvere queste problematiche partendo dai principi generali dell'ordinamento e utilizzando, per ora, la normativa vigente e delineando al contempo un quadro regolatorio che possa, in futuro, disciplinare l'IA utilizzando un approccio non contenitivo, punitivo o repressivo, bensì flessibile e agevolativo, ferma restando la necessità di approntare i necessari presidi giuridici a tutela dei diritti fondamentali della persona umana (e delle imprese interessate dagli effetti dell'IA).

Il tema dell'intelligenza artificiale deve essere affrontato sia in una prospettiva *de iure condito* (alla luce della normativa vigente e dei principi generali anche europei in quanto l'IA è un fenomeno rilevante nel quadro multilivello essendo fondamentale l'apporto dell'Unione europea), che in una visione *de iure condendo* (quindi tenendo conto anche dei testi normativi che formano oggetto di negoziati a livello di Unione europea e del recente Disegno di Legge governativo).

2. Intelligenza artificiale e umana, input e output, strumentalità e riserva di umanità

Si parla di *intelligenza artificiale* in quanto trattasi di una forma di intelligenza ricollegata ad una macchina e, quindi, non umana.

A tal uopo è necessario rievocare alla memoria il famoso *test* elaborato negli anni '50 da uno celebre scienziato, il quale sostenne che si sarebbe potuto assegnare l'aggettivo di "intelligente" ad una macchina solo nel caso in cui l'*output* generato da quest'ultima fosse stato indistinguibile da quello formato da un essere umano, profilandosi così, per la prima volta nella storia, il requisito della impossibilità/difficoltà di distinguere l'*output elettronico* da quello *umano*.

Ove, quindi, la macchina riesca a dare un risultato assolutamente simile ed indistinguibile da quello che fornisce l'intelligenza umana si è al cospetto di "macchina intelligente" ovverosia di un'intelligenza artificiale (IA): questo è il c.d. "test di *Turing*".

Oggi, dopo oltre 70 anni da quelle elaborazioni teoriche, si inizia finalmente a "vedere" e "toccare con mano" l'intelligenza artificiale. Ciò grazie alle imponenti capacità di calcolo delle macchine (soprattutto nell'elaborazione e nello *storing* dei dati), capacità che, un tempo, non era inconfigurabile e inimmaginabile dal punto di vista tecnico. Quindi, con riguardo all'intelligenza artificiale (atteso che oramai del tutto legittimamente si può discorrere di *intelligenza artificiale* in senso tecnico poiché la macchina effettivamente elabora dati come fosse un uomo e, anzi, talvolta meglio di un essere umano), i problemi si collocano sia a monte che a valle.

A monte i problemi sorgono poiché è necessario immettere dati e, quindi, fornire alla macchina il c.d. *input* (quel che viene immagazzinato dalla macchina). Qui sorgono soprattutto questioni correlate alla tutela del diritto di autore, a quella delle banche dati e, in generale, alla tutela della conoscenza umana e degli strumenti attraverso cui essa viene tramandata. Con queste macchine e con le IA accade che, in blocco, una conoscenza (che un essere umano o più esseri umani faticano una vita ad acquisire), riesca ad essere incamerata dalla macchina immediatamente o, comunque, in tempi molto più rapidi e senza particolari sforzi da parte dell'IA. Si profila altresì il problema dell'utilizzo di quei dati che, spesso, sono presenti in archivi giornalistici e in banche dati pubbliche o private (a fatica formata da una generalità di persone con un'attività pluriennale). Vi sono, quindi, numerose questioni teorico-pratiche che si collocano a monte nella fase di acquisizione e lavorazione dei dati costituenti l'*input*.

A valle i problemi investono l'*output* ovverosia ciò che la macchina "crea" attraverso complessi meccanismi di assemblaggio e rimescolamento dei dati originati dall'*input* immesso. In tale fase la macchina rielabora tutti i dati immessi e ci si è chiesti se sia possibile parlare di un "prodotto" diverso. Inoltre, c'è il tema dell'intelligenza artificiale auto-generativa (o generativa), ossia una intelligenza artificiale che, come un essere umano, identifica il bisogno di conoscenza (soddisfacendolo con l'acquisizione di nuovi dati rinvenuti on line) o l'esigenza di rielaborazione dei dati immessi al fine di creare nuove elaborazioni. Trattasi di una macchina (o IA) che si sviluppa autonomamente secondo l'algoritmo o persino modificando autonomamente l'algoritmo originario fissato dall'elaboratore del *software*. L'IA generativa diventa così una "macchina" che sfugge al controllo persino di coloro che hanno progettato il *software*, atteso che quest'ultimo viene cambiato, modificato e aggiornato dalla macchina stessa.

Senza entrare nel campo dei linguaggi tecnici di programmazione, "ingegneri e architetti informatici" evidenziano che vi sono nuovi meccanismi di elaborazione utilizzati dall'IA che consentono alla macchina di auto-modificarsi, di identificare in autonomia l'esigenza di conoscenza e di provvedere a quella esigenza di conoscenza, fornendo nuovi *output*. Esistono persino macchine che generano *software* in base ad una specifica richiesta o sulla base di esigenze generative poste dalla medesima IA.

Quindi, con riguardo soprattutto a quest'ultima intelligenza (*generative IA*), si pone un problema molto rilevante, anche di natura etica e sociale (si pensi, ad esempio, alla salvaguardia dei posti di lavoro). Ci si può chiedere se il giudice in carne ed ossa possa essere sostituito dall'intelligenza artificiale (si pensi al c.d. giudice *Robot*). In relazione al funzionario pubblico si è avuta già realizzata un'ampia spersonalizzazione: il responsabile del procedimento, come diceva un autorevole giurista, oramai è sovente divenuto, almeno in relazione ad alcune attività, una casella di posta elettronica a risposta automatizzata e oggi la tecnologia consente che si trasformi in un *robot* o in un algoritmo. Queste sono le ragioni per cui si è emersa, a livello globale, la necessità di salvaguardare la c.d. "riserva di umanità", questione risolta affermando il principio che a livello regolatorio debbano essere fissate alcune limitazioni (come ha fatto il recentissimo disegno di legge governativo) anche al fine di assicurare l'operatività delle responsabilità e dei controlli, nonché la tutela delle situazioni soggettive coinvolte dall'esercizio dei poteri amministrativi e giudiziari attraverso strumenti di IA.

Le problematiche, tuttavia, investono anche il mondo privato ed in particolare l'area delle professioni intellettuali. Nella evoluzione della civiltà le macchine hanno da tempo sostituito larghe aree di lavoro manuale (eliminando anche tanti antichi mestieri). Oggi le macchine (e i loro ideatori) puntano a sostituire i professionisti intellettuali (quindi il cervello, apparato che ha sempre contraddistinto l'essere umano). Si pensi alle macchine (IA) che elaborano, in automatico, contratti o atti defensionali oppure capaci di prevedere il probabile esito di una controversia (già esistono IA in grado di stipulare contratti di transazione, contratti di lavoro). Si pensi altresì alle IA di progettazione di interni.

Dunque, in pratica, la macchina, pian piano, sta sostituendo l'uomo e persino il professionista intellettuale.

Oltre a questi problemi, che hanno ineluttabili ricadute sulle relazioni sociali, sussiste, inoltre, il tema dell'utilizzo della macchina (modalità e ambiti). Si pensi, ad esempio, all'utilizzo dell'intelligenza artificiale nel settore sanitario per le necessità diagnostiche e, addirittura, terapeutiche. Già è possibile operare a distanza attraverso le nuove tecnologie. L'intelligenza artificiale viene anche usata, in campo medico, per diagnosticare malattie e, in alcuni casi, per condurre anche alcuni interventi. Ci si può chiedere sino a che punto la macchina possa sostituire il sanitario. Sul punto interviene il recente Disegno di Legge.

Emergono anche altre rilevanti tematiche. Si pensi al controllo biometrico da remoto, in diretta o in differita. Si immaginino quei sistemi dove la macchina, attraverso telecamere di videosorveglianza, controlli la popolazione. Ci si può chiedere sino a che punto possa essere utilizzata la macchina a fini di controllo, prevenzione e contrasto delle attività illecite. Emerge, difatti, tutto il tema della privacy. Sul tema interviene il legislatore europeo, il quale fissa dei limiti a tutela proprio dei diritti fondamentali della persona, prevenendo l'utilizzo di sistemi biometrici di controllo in via eccezionale ed esclusivamente per contrastare alcuni reati elencati da un atto dell'Unione europea.

Prendendo atto che l'intelligenza artificiale solleva tantissimi problemi, passiamo ad esaminare le principali questioni giuridiche e le soluzioni sinora proposte.

3. La c.d. "personalità giuridica elettronica" e il c.d. "Robot-schiavo": l'IA resta una res

La questione della "personalità giuridica elettronica" è stata sollevata persino nell'ambito dei negoziati a livello europeo. In una Risoluzione del Parlamento europeo è sollecitata una riflessione scientifica sulle tecniche giuridiche idonee ad attribuire una personalità giuridica a *robot*, macchine e IA.

In proposito la scienza giuridica europea ha elaborato una serie di teorie attingendo all'esperienza dei romani (ad esempio invocando il regime giuridico applicabile agli atti degli schiavi).

L'IA (o un robot o una macchina) è pur sempre una *res* che può formare solo *oggetto* di rapporti giuridici, escludendosi che possa assurgere a *soggetto* di diritti o centro di imputazione di rapporti giuridici.

La dottrina prevalente, perorando la tesi contraria all'attribuzione della personalità giuridica alla macchina, ha proprio proposto il modello della "macchina-schiavo".

La questione molto simile a quella di recente posta in relazione agli animali. Sono note le più recenti teorie (persino riferite alla posizione degli animali in giudizio) sulla natura giuridica dell'animale che, nonostante le nuove tutele sviluppate dall'ordinamento, resta pur sempre una *res* escludendosi che possa divenire soggetto di diritto. La tesi attualmente prevalente ritiene che l'animale, anche se ampiamente tutelato (anche dalla disciplina penale), continui ad essere una *res*, seppure *senziente*.

Anche per l'intelligenza artificiale (e, dunque, per il *robot* o per la macchina), anche alla luce dei già menzionati rapporto di strumentalità e principio di riserva di umanità, l'orientamento prevalente ritiene che non sia possibile identificare una soggettività giuridica elettronica, riferibile all'IA, restando quest'ultima una *res* senza che sia ipotizzabile un'IA (anche se generativa) identificabile quale centro di imputazione di rapporti giuridici.

La macchina e, quindi, l'intelligenza artificiale, – osserva tale tesi – non può essere utilizzata come schermo (e dunque quale nascondiglio protettivo) da chi intenda, attraverso la macchina, cagionare danni a terzi e, in tal modo, realizzare illeciti civili e illeciti penali. Il carattere generativo di alcune IA non può quindi essere invocato per recidere il nesso di causalità (o bloccare imputazioni colpose o dolose) da parte di chi abbia progettato, ideato, utilizzato la macchina oppure da chi ne sia proprietario.

Il legislatore italiano, con il menzionato disegno di legge governativo, dà un messaggio molto chiaro da questo punto di vista, in perfetta sintonia, peraltro, con quanto emerge dai testi in discussione a livello europeo.

Le regole della responsabilità (civile e penale), come noto, hanno natura di ordine pubblico e non possono essere eluse. Il Legislatore, peraltro, è pienamente titolato a disciplinare il regime giuridico dei danni derivanti dagli *outputs* dell'intelligenza artificiale. In un quadro di attento bilanciamento di interessi le linee di politica legislativa non intendono determinare blocchi allo sviluppo delle nuove tecnologie, soltanto introdurre limitazioni funzionali alla tutela dei diritti fondamentali. In questa prospettiva l'operatività delle regole della responsabilità, da riferirsi pur sempre ad una persona fisica o ad una persona giuridica, deve essere sempre garantita al fine di assicurare una piena tutela giuridica.

Il problema così diviene quello della tecnica di imputazione dell'atto (lecito o illecito) compiuto dalla macchina e del regime di responsabilità applicabile alla specie.

4. L'imputazione all'IA di atti favorevoli. L'analisi delle controversie in materia di tutela della proprietà intellettuale e industriale. L'impossibilità di attribuire direttamente all'IA la qualifica di inventore e/o autore

Come anticipato, con riguardo alla imputabilità dell'azione della macchina, favorevole o sfavorevole, si può analizzare, innanzitutto, il tema degli atti favorevoli in relazione ai quali si è fuori dell'area delle

problematiche connesse ad eventuali responsabilità civili o penali. Il tema si è posto nell'area della tutela della proprietà intellettuale e della proprietà industriale.

A livello globale le principali giurisdizioni nazionali sono state interessate da domande di brevetto direttamente riferite all'IA quale inventore: il richiedente evidenziava che era stata proprio la macchina, con la sua intelligenza artificiale generativa, previa autonoma identificazione dell'esigenza di utilità sociale giustificante l'attività inventiva, ad ideare nuove tecniche e prodotti, ad impiego industriale, fino a quel a quel momento inesistenti, idonei a risolvere innovativamente quelle determinate problematiche. Di qui le richieste di brevetto dell'invenzione elaborata dalla intelligenza artificiale, direttamente a nome dell'intelligenza artificiale stessa presentate in varie giurisdizioni (Stati Uniti, Europa, Australia, Sud Africa, etc.).

La risposta, quasi unanime, è stata negativa.

A livello globale, difatti, nei principali ordinamenti giuridici esiste il principio (normativamente sancito o pretoriamente affermato) della riferibilità della richiesta di brevetto alla persona fisica che abbia inventato quel nuovo *quid* utile all'industria (al più è possibile l'imputazione ad una persona giuridica). In sostanza deve esistere un centro di interessi idoneo a costituire soggetto di rapporti giuridici quantomeno patrimoniali (il titolare del brevetto deve avere una personalità giuridica). Per l'effetto in quasi tutte le giurisdizioni le richieste di brevetto direttamente riferite all'IA generativa sono state rigettate (solo in pochi Paesi tale richiesta è stata accolta poiché la normativa ivi esistente non prevede un vincolo di riferibilità dell'invenzione ad una persona fisica o giuridica).

Accanto alla tutela riconnessa alla proprietà industriale si è posto il problema della tutela autoriale.

Ci si è chiesti se sia possibile identificare l'intelligenza artificiale come autore di un'opera dell'ingegno.

Anche sul punto, generalmente, la risposta è stata negativa, in quanto, in base ai principi normativi della legislazione autoriale (peraltro molto simili in tutti gli Stati a livello globale), la tutela del diritto d'autore è ricollegata ad un atto umano ideativo e creativo avente i caratteri dell'originalità (c.d. requisito dell'atto umano originariamente creativo) che dia luogo ad un'opera diversa da altra esistente. È necessaria, quindi, una attività ideativa, creativa e di esteriorizzazione, che può essere tanto al pubblico, quanto ad una cerchia limitata ed autorizzata di soggetti.

Dal punto di vista del *copyright* e, dunque, della c.d. proprietà intellettuale, la tutela direttamente imputata all'intelligenza artificiale è stata esclusa, in quanto deve esserci sempre un riferimento all'azione dell'essere umano.

Escludendosi, quindi, che l'IA possa essa stessa direttamente essere considerata "autore" di un'opera protetta con gli strumenti della tutela autoriale, il problema è stato posto da un altro angolo prospettico (la sollecitazione è stata colta dalla giurisprudenza di legittimità che ha formulato condivisibili principi).

Cosa accade ove l'autore, per creare un'opera originale ed innovativa, utilizzi la macchina (IA)? Ci si è chiesti se questo atto creativo, rispetto al quale l'essere umano ottenga ausilio dalla macchina attraverso il suo utilizzo, possa essere tutelato dalla disciplina sul diritto d'autore.

Si pensi alla *digital o computer art* (arte digitale computerizzata o forma d'arte riconnessa all'elettronica): una fotografia elaborata (o rielaborata) da un essere umano ma attraverso l'ausilio dell'intelligenza artificiale è un *quid* tutelabile dalla legislazione in materia di diritto d'autore.

In tale ipotesi la giurisprudenza ha dato una risposta molto chiara e che, peraltro, riecheggia nel disegno di legge governativo sull'IA. Si è sostenuto che *bisogna verificare l'esistenza del contributo umano, verificando cosa abbia fatto l'essere umano e che tipo di contributo abbia dato alla macchina, al fine di identificare una creazione originale riconducibile all'uomo (sia pure attraverso l'utilizzo della macchina). Colui che invochi la tutela autorale deve dare la prova dell'essere effettivamente il creatore di quell'output umano, anche se realizzato usando l'intelligenza artificiale. Quindi, secondo la Corte di cassazione, è necessario effettuare una indagine in concreto avente ad oggetto l'apporto umano a fine di capire se l'azione della macchina abbia o meno avuto una natura prevalente rispetto al contributo dell'uomo.*

Tali principi sembrano emergere dal recente disegno di legge governativo in materia di AI: "possono essere tutelate dal diritto di autore le opere di cui ai numeri precedenti, create con l'ausilio di algoritmi di intelligenza artificiale, nel caso in cui il contributo umano nella ideazione e realizzazione dell'opera, tramite l'uso dell'algoritmo, sia creativo, rilevante e dimostrabile, senza pregiudizio dei diritti del creatore del programma per elaboratore, attraverso il quale è implementato l'algoritmo". Trattasi di una disposizione che, a dispetto della semplicità e della condivisibilità degli assunti di base, pone all'interprete alcune problematiche. Lasciando al giudice la risoluzione del problema, pone, da un lato, il delicatissimo compito di identificare "quanto" sia riferibile alla macchina e "quanto" sia riferibile all'autore dell'opera e, dall'altro, il compito di identificare i *diritti del creatore del programma*. La norma presuppone (senza tuttavia identificarli e definirli) l'esistenza di diritti anche in favore di colui che abbia programmato l'IA in quanto essa impone espressamente la salvaguardia dei *diritti del creatore del programma* nella parte in cui afferma "senza pregiudizio per i diritti del creatore del programma per elaboratore, attraverso il quale è implementato l'algoritmo".

Dunque, la norma sta facendo riferimento a “*diritti*” (attribuiti all’elaboratore del *software* ossia all’elaboratore dell’intelligenza artificiale) che non vengono identificati né qualificati. Al contempo, non si nega l’esistenza del diritto di autore in capo all’essere umano che abbia utilizzato l’intelligenza artificiale, purché vi sia un contributo creativo rilevante e dimostrabile. Sembra che la quota parte riconducibile all’IA debba essere riferita al creatore del *software*. Tuttavia, cosa accade se sia stata direttamente l’IA a generarlo o se il creatore del *software* non coincida con il proprietario dell’IA?

Concludendo sul punto, anche l’indagine condotta nell’area della tutela della proprietà intellettuale e industriale, conferma che non è possibile attribuire all’IA una personalità giuridica elettronica. Le linee di politica legislativa lo escludono, anche in una prospettiva *de iure condendo*, in quanto, ad esempio, il richiamato disegno di legge governativo riferisce *i diritti*, al più, a colui che abbia elaborato il *software* e non all’IA.

Tutte le giurisdizioni hanno escluso l’attribuzione della tutela prevista in materia di proprietà industriale e intellettuale direttamente alla macchina. Ogni rapporto giuridico deve essere sempre riferibile ad una persona fisica o ad una persona giuridica, ossia a quei centri di imputazione che l’ordinamento giuridico riconosce. La macchina e l’intelligenza artificiale (anche se generativa), secondo gli orientamenti prevalenti, restano un oggetto (*res*) a servizio dell’essere umano.

5. *IA e profili problematici di responsabilità civile: il caso dei veicoli a guida autonoma. L’applicazione dei criteri di imputazione previsti dal Codice civile e i rapporti con la disciplina eurounitaria approvata e approvanda*

Dopo aver descritto cosa sia l’intelligenza artificiale e, soprattutto, dopo aver affrontato il tema, logicamente preliminare, delle imputazioni e della personalità giuridica elettronica, ci si deve confrontare con le problematiche specifiche sorte nell’applicazione delle regole in materia di responsabilità civile. Dall’esame degli studi italiani e stranieri emerge che il problema si è posto soprattutto con riguardo all’utilizzo di veicoli a guida autonoma.

In proposito deve preliminarmente osservarsi che i veicoli, per circolare in un determinato territorio, devono aver ottenuto l’immatricolazione secondo la disciplina ivi vigente. Nel contesto europeo esiste una regolazione completamente diversa da quella esistente in altri Stati (ad esempio dagli Stati Uniti): in Europa, difatti, non possono circolare veicoli a guida completamente autonoma, ammettendosi una guida meramente “assistita”, con svariati livelli di autonomia. In Europa, quindi (persino in Germania che ad oggi possiede l’ordinamento giuridico più avanzato quanto a livello di autonomia ammissibile), non si raggiungono quei livelli di totale autonomia di guida presenti (e legittimi) negli Stati Uniti, giurisdizione in cui, more solito, ha dato la scaturigine alla prima casistica riconnessa allo sviluppo di tale tecnologia. Ci si riferisce alle controversie in materia di incidenti determinati da veicoli a guida completamente autonoma, in relazione ai quali si sono registrate condanne per omicidi colposi a carico di soggetti cui è stato addebitato, appunto, di non aver colposamente ripreso il controllo del veicolo bloccando la funzione di guida assistita. Il tema della “guida autonoma” e, dunque, quello della responsabilità derivante dall’*output* da intelligenza artificiale che consente la “guida autonoma”, risiede proprio nello sviluppo di una tecnologia che, da un lato, consente al veicolo di accelerare e decelerare, sterzare, parcheggiare tenendo conto della posizione nello spazio anche con riguardo alla strada, agli altri veicoli, ai pedoni, agli eventuali ostacoli e alla destinazione predeterminata, anche prevenendo incidenti in maniera indipendente dal controllo umano, ma, dall’altro, permette all’uomo alla guida di riprendere il controllo del veicolo laddove non vi siano più quei presupposti per consentire una guida automatica sicura.

In Italia la normativa vigente consente solo un livello non avanzato di guida autonoma con impossibilità di immatricolare veicoli a “guida totalmente autonoma”: si è in presenza di un livello di autonomia che impone al guidatore di rimanere sempre al vigile controllo del veicolo con necessità di riprendere i comandi laddove non vi siano più le condizioni per portare avanti la “guida assistita”.

Dal punto di vista dei criteri di imputazione della responsabilità in questo settore specifico la dottrina ha segnalato la necessità di applicare i criteri generali previsti dall’art. 2054 c.c.

In altre giurisdizioni (in Germania e in Francia), invece, il Legislatore è intervenuto identificando specifici criteri di imputazione alle fattispecie di responsabilità da danni cagionati da veicoli a guida autonoma.

In Italia la dottrina ha posto il problema relativo al fatto che il 2054 c.c. potrebbe non essere una norma aggiornata, in quanto tale articolo lascerebbe sul tappeto il problema della responsabilità del produttore del veicolo. Esso prende a riferimento la responsabilità del proprietario e quella del conducente, ma ci si chiede come debba configurarsi la responsabilità del produttore del veicolo. Atteso che l’intelligenza artificiale, in questa ipotesi, opera attraverso un prodotto (un’autovettura dotata di *software*), dovrebbe applicarsi anche la disciplina di derivazione eurounitaria relativa ai danni da prodotti difettosi, dimostrando, ovviamente, l’esistenza dei relativi elementi costitutivi (difetto, danno e del nesso di causalità tra difetto e danno).

Altri autori, invece, hanno invocato l'applicazione anche di altre norme.

La problematica relativa alla responsabilità da IA (non solo quella connessa alla guida autonoma dei veicoli) si allarga così alla verifica dell'applicazione delle norme generali del Codice civile. Distaccando lo sguardo dalla fattispecie specifica riguardante i veicoli a guida autonoma (dove generalmente si fa riferimento all'art. 2054 e alla disciplina in materia di responsabilità da prodotti), allargando il campo di osservazione, il tema della responsabilità per danni cagionati attraverso l'utilizzo dell'IA può riferirsi a tutta l'attività dell'intelligenza artificiale e, dunque, in generale ai pregiudizi determinati dall'intelligenza artificiale. Su tale presupposto, pertanto, altra parte della dottrina ha invocato l'applicazione degli altri criteri speciali previsti in materia di responsabilità extracontrattuale.

Si è fatto riferimento, ad esempio, all'art. 2051, ossia al criterio che stabilisce la responsabilità per cose in custodia. Oppure, sempre con riguardo alla guida di veicoli con guida autonoma, ossia con una guida fondata su un sistema di intelligenza artificiale, si è invocata anche l'applicazione del 2050 c.c., considerando l'attività di guida autonoma una attività pericolosa.

Con riguardo al discorso più generale dei danni da utilizzo dell'IA, c'è, quindi, chi ha invocato anche il criterio di imputazione di cui al 2051 (ossia la responsabilità per cose in custodia), chi ha fatto riferimento all'art. 2049 c.c. e all'art. 1228 c.c., considerando l'intelligenza artificiale, quindi, il *robot-schiavo* (secondo l'immagine poc'anzi ricordata), una specie di *preposto*, *commesso* o *ausiliario*.

Questo è in sintesi il dibattito che si sta svolgendo *de iure condito*: si discute su quelli che potrebbero essere i criteri di imputazione della responsabilità civile in relazione a danni prodotti dalla intelligenza artificiale (o attraverso l'utilizzo della intelligenza artificiale).

La dottrina si è posta anche altri problemi.

I danni cagionati dall'utilizzo di veicoli dotati di sistemi di guida autonomi possono plasticamente essere ricondotti alla responsabilità da prodotto. In tal caso l'intelligenza artificiale (la guida autonoma dell'IA) è permessa in quanto la macchina (l'autovettura) possiede un *software* che valuta la strada, la presenza di pedoni o di veicoli, l'andamento dell'autovettura, la corsia e, dunque, in generale, la collocazione spaziale del veicolo, identificando le azioni da compiere, sostituendosi, in definitiva, alla guida umana tradizionale. In tale fattispecie pare proprio esservi un servizio riconnesso al prodotto, evenienza che consente l'applicazione della Direttiva sulla responsabilità del produttore.

Il problema, però, si pone quando l'operatività dell'intelligenza artificiale sia svincolata da un prodotto specifico.

Si immaginino le ipotesi in cui l'intelligenza artificiale operi quale mero *software*. In tali casi ci si avvicina ai servizi digitali emergendo così il problema dell'applicabilità della Direttiva relativa alla responsabilità del produttore.

È per questo, dunque, che il Legislatore europeo sta intervenendo in materia ed è proprio per tale ragione che, da anni, si negoziano atti di diritto eurounitario che si vuol far transitare dall'area della *soft law* a quella della *hard law*. È noto che è allo studio una proposta di Regolamento sulla AI, oramai "alle porte" essendo stata da poco approvata in via definitiva dalle Istituzioni EU, e vi è anche una proposta di modifica della direttiva sui prodotti difettosi, nonché una proposta di direttiva specifica volta ad introdurre norme sulla responsabilità da *output* della IA.

Prima di esaminare il disegno di legge governativo italiano è necessario, quindi, esaminare questi atti di diritto eurounitario.

6. Le linee di politica legislativa europea: l'EU IA Act 2024

Il legislatore comunitario intende bilanciare i due principali interessi in campo, consentendo lo sviluppo delle nuove tecnologie connesse all'IA e, al contempo, tutelando i diritti fondamentali delle persone e i principali interessi generali toccati dall'operatività dell'IA.

Si pensi, ad esempio, a tutto il contenzioso – peraltro non solo italiano, ma anche straniero – che ha evidenziato come l'IA possa dare luogo a rilevanti disparità di trattamento. L'algoritmo, infatti, potrebbe essere "pensato male", ossia impostato erroneamente, determinando *output* discriminatori, ad esempio, per ragioni di salute, sessuali, politiche, religiose o socio-economiche. In Italia la giurisprudenza si è occupata dell'algoritmo discriminatorio governante procedure di assunzione o di scorrimento di graduatorie (con riguardo ai docenti, ad esempio, era stato creato un algoritmo volto all'assegnazione automatica delle sedi di servizio che aveva prodotto esiti illogici, irragionevoli e discriminatori, proprio in quanto perché era stato impostato in modo erroneo). Oltre al problema della sostituzione della macchina all'uomo in ambito lavorativo, una delle più grandi preoccupazioni investe l'azione della macchina ovvero la ipotesi in cui l'IA venga usata, con esiti discriminatori, quale strumento a supporto dell'azione della pubblica amministrazione o di privati.

Si profila altresì il tema del controllo generalizzato della cittadinanza attraverso tecniche di riconoscimento biometrico a distanza che possono essere utilizzate nell'interesse pubblico, da bilanciarsi, tuttavia, con quello alla privacy (per tale ragione dovrebbe trattarsi di ipotesi eccezionali).

Su tali tematiche interviene l'IA Act europeo.

Il regolamento europeo fornisce una concettualizzazione dell'intelligenza artificiale: *“con il termine intelligenza artificiale si indica una famiglia di tecnologie, in rapida evoluzione, in grado di apportare una vasta gamma di benefici economici e sociali in tutto lo spettro delle attività industriali e sociali. L'uso dell'intelligenza artificiale, garantendo un miglioramento delle previsioni, l'ottimizzazione delle operazioni e dell'assegnazione delle risorse e la personalizzazione dell'erogazione dei servizi, può contribuire al conseguimento di risultati vantaggiosi dal punto di vista sociale e ambientale, nonché fornire vantaggi competitivi fondamentali alle imprese e all'economia europea”*.

IA Act europeo da conto anche dell'esistenza di rischi: *“gli stessi elementi e le stesse tecnologie che alimentano i benefici socio-economici dell'IA possono, altresì, comportare nuovi rischi o conseguenze negative per le persone giuridiche e la società”* e, allora, *“l'Unione europea si impegna a perseguire un approccio equilibrato. L'interesse dell'Unione è quello di preservare la leadership tecnologica dell'UE e assicurare che i cittadini europei possano beneficiare delle nuove tecnologie sviluppate operanti in conformità ai valori, ai diritti fondamentali e ai principi dell'Unione”*.

Si fissa, così l'obiettivo di sviluppare una IA affidabile, sicura, conforme alla Legge e a tutela dei diritti fondamentali.

Per quanto concerne la tutela dei diritti fondamentali, nell'IA Act si dice che esso, attraverso la definizione di requisiti per una IA affidabile e di obblighi proporzionati per tutti i partecipanti alla catena del valore, migliorerà e promuoverà la protezione dei diritti tutelati dalla Carta (facendosi riferimento alla protezione della vita privata, dei dati personali, della dignità umana, della parità tra donne e uomini; si afferma la necessità di prevenire un effetto dissuasivo sui diritti di libertà di espressione e di riunione; si esplicitano numerose previsioni relative a rischi di decisioni assistite dalle IA errate o distorte in settori critici quali l'istruzione, la formazione, l'occupazione e le attività di contrasto di reati, dovendosi, in tutti questi casi, implementarsi regole idonee a prevenire ed evitare i predetti rischi).

Altra esigenza che deve essere soddisfatta è della prevenzione dell'utilizzo di tecniche subliminali che puntino a pratiche manipolative o di sfruttamento dell'essere umano. Emerge il problema del marketing, di ciò che viene offerto all'essere umano, ossia il tema dell'essere umano “guidato” dall'algoritmo e dall'IA nei propri desideri e negli atti di consumo. E, dunque, necessario prevenire e contrastare queste tecniche di manipolazione della collettività e dell'individuo.

Circa il come si debbano raggiungere questi obiettivi, con lo *European Union IA Act*, ossia il nuovo regolamento eurounitario in materia (atto di *hard law* con effetti vincolanti che soppianta i tradizionali atti di *soft law*: risoluzioni, comunicazioni, libri bianchi etc.), la logica europea è, innanzitutto, quella di considerare l'intelligenza artificiale un prodotto da sottoporre alla disciplina generale sulla sicurezza.

È così consentito sviluppare le tecnologie basate sull'IA ma il prodotto risultante deve essere sicuro e non difettoso.

Il quadro giuridico di riferimento fa capire che l'Unione europea introduce un sistema di controllo sul prodotto *intelligenza artificiale*, identificando le c.d. “intelligenze artificiali” *ad alto rischio*. In relazione a queste tipologie di IA (cioè, quelle idonee a determinare effetti negativi sui diritti fondamentali delle persone), l'Unione europea introduce dei vincoli di controllo, *ex ante* e *ex post*, dunque, un sistema di certificazione di qualità come avviene, appunto, per i prodotti. Sono introdotti vincoli operativi intrinseci e estrinseci. Nel regolamento europeo, quindi, l'IA è considerata un vero e proprio prodotto.

6.1. (Segue) *L'adeguamento della direttiva sulla responsabilità del produttore alle problematiche poste dall'IA: il rapporto tra disciplina europea sulla responsabilità del produttore, altre normative eurounitarie specifiche e discipline nazionali contrattuali e extracontrattuali*

Il quadro regolatorio europeo viene completato con una proposta di modifica della Direttiva in materia di responsabilità del produttore, che si prefigge una serie di obiettivi tra cui:

- *l'estensione del concetto di prodotto (in modo che possa essere riferito anche all'intelligenza artificiale);*

- *l'alleviamento dell'onere della prova (trattandosi -come dice il legislatore europeo- di fattispecie complesse dove è difficile, spesso, identificare quel nesso di causalità tra danno e difetto nella specie proveniente dal funzionamento dell'intelligenza artificiale).*

Vi è poi un problema di asimmetria informativa che il Legislatore europeo intende risolvere, in primo luogo, modificando la Direttiva in materia di responsabilità del produttore e, in secondo luogo, introducendo norme specifiche in materia di responsabilità da intelligenza artificiale che andranno ad integrare, secondo l'intenzione del legislatore comunitario, i sistemi nazionali in materia di responsabilità civile.

Si deve ricordare che la Direttiva in materia di responsabilità del produttore prevede, secondo l'idea eurounitaria (ossia *l'idea della Corte di giustizia e della Commissione europea*), una responsabilità di tipo oggettivo (fondata sul mero nesso di causalità) ed esclusivo (in quanto, ove si applichi la Direttiva sulla responsabilità del produttore non devono trovare applicazione le norme interne che prevedano criteri di imputazione di responsabilità di tipo meramente oggettivo, restando operativi i criteri nazionali fondati sulla colpa, una colpa, secondo alcuni, estesa anche alla colpa c.d. presunta ma con possibilità di fornire la prova contraria).

Nella proposta di modifica della direttiva sulla responsabilità del produttore si richiamano gli esiti cui è giunta la Corte di giustizia con riguardo ai rapporti tra regime eurounitario in materia di responsabilità del produttore e regimi nazionali sulla responsabilità contrattuale: *il sistema della responsabilità del produttore esclude i criteri di imputazione nazionale fondati sul mero nesso di causalità, ma concorre con la disciplina prevista, per esempio, in materia di responsabilità contrattuale.*

Circa i rapporti con le altre discipline europee che governano specifici casi di responsabilità civile si specifica che, a livello UE, esistono numerosi *strumenti complementari in materia di responsabilità*, facendo riferimento alla disciplina in materia di vendita di beni di consumo, di fornitura di servizi e prodotti digitali, di responsabilità ambientale o da illecito trattamento di dati personali. Quindi il blocco normativo sulla responsabilità del produttore concorre con queste regole specifiche (le c.d. "altre responsabilità europee") che prevedono diversi criteri di imputazione della responsabilità.

Fatta questa premessa, che investe il campo di applicazione della disciplina in materia di responsabilità del produttore, è necessario esaminare gli obiettivi della proposta e gli strumenti pensati per perseguirli.

Innanzitutto, come anticipato, essa intende allargare la *nozione di prodotto* essendo allo stato dubbia l'applicazione della disciplina in materia di responsabilità del produttore ai danni determinati dall'utilizzo della intelligenza artificiale, specie nelle ipotesi in cui l'IA non fosse collegata ad un *quid* qualificabile in senso tecnico come prodotto. Il nuovo testo proposto chiarisce che *i sistemi di intelligenza artificiale e i beni basati sull'intelligenza artificiale sono prodotti e, pertanto, rientrano nell'ambito dell'applicazione della Direttiva sulla responsabilità per il danno da prodotti difettosi, il che implica, come qualsiasi altro prodotto, che il danneggiato possa chiedere il risarcimento del danno causato dal sistema di intelligenza artificiale difettoso, senza dover provare la colpa del fabbricante.*

In secondo luogo, la proposta prevede espressamente che *possono essere ritenuti responsabili, oltre ai fabbricanti di hardware, anche i fornitori di software e di servizi digitali, che incidono sul funzionamento del prodotto, come nel caso del servizio di navigazione in un veicolo autonomo.* Vi è una chiara identificazione dei soggetti responsabili ai sensi della Direttiva in questione. La proposta, inoltre, *garantisce che i fabbricanti possano essere ritenuti responsabili per le modifiche da essi apportate a prodotti che hanno già immesso sul mercato, anche quando tali modifiche sono dovute ad aggiornamenti del software o ad apprendimento automatico.*

In terzo luogo, *si alleggerisce l'onere della prova nei casi complessi (tra cui possono figurare determinati casi relativi a sistemi di IA) e nei casi in cui i prodotti non sono conformi ai requisiti di sicurezza.*

Posti questi tre obiettivi generali il Legislatore europeo precisa che *il prodotto è considerato difettoso se non offre la sicurezza che il grande pubblico può legittimamente attendersi, tenuto conto di tutte le circostanze.* Si definisce il prodotto difettoso e, a monte, il "prodotto" medesimo come ogni bene mobile, anche se integrato in un altro bene mobile o in un bene immobile, inclusi elettricità, file per la fabbricazione digitale e software (il file per la fabbricazione digitale è una versione o un modello digitale di un bene mobile). Dunque, la nozione di prodotto viene estesa anche all'intelligenza artificiale e, quindi, al software su cui si fonda l'intelligenza artificiale.

Tra gli aspetti più importanti emerge la rimodulazione dell'onere della prova. Nel solco della tradizione si ribadisce che *gli Stati membri provvedano affinché l'attore sia tenuto a provare il carattere difettoso del prodotto, il danno subito e il nesso di causalità tra il difetto e il danno.* In relazione all'intelligenza artificiale, come per altre fattispecie critiche (si pensi al danno da farmaci), ossia quelle in cui sia complesso riuscire a fornire e a soddisfare l'onere della prova relativo al nesso di causalità tra difetto e danno, il Legislatore comunitario intende alleviare l'onere della prova in capo all'attore introducendo una *presunzione di difettosità*: *"si presume il carattere difettoso di un prodotto qualora sia soddisfatta una delle seguenti condizioni:*

- *il convenuto non ha rispettato l'obbligo di divulgare gli elementi di prova pertinenti;*
- *l'attore prova che il prodotto non rispetta i requisiti di sicurezza;*
- *l'attore prova che il danno è stato causato da un malfunzionamento evidente del prodotto durante l'utilizzo normale o in circostanze ordinarie.*

Si presume l'esistenza del nesso di causalità tra il difetto del prodotto e il danno nel caso in cui sia stato provato che il prodotto sia difettoso e che la natura del danno cagionato sia generalmente coerente con il danno in questione, nonché qualora l'Organo giurisdizionale nazionale rilevi che l'attore incontra difficoltà

eccessive a causa della complessità tecnica e scientifica nel provare il carattere difettoso del prodotto, il nesso di causalità e il danno (ed è questo il caso dell'intelligenza artificiale).

Si presumono il carattere difettoso del prodotto e il nesso di causalità tra difetto e danno, o entrambi tali elementi, se l'attore ha dimostrato, sulla base di elementi di prova sufficientemente pertinenti, che il prodotto ha contribuito a cagionare il danno ed è probabile che il prodotto fosse difettoso, oppure che il carattere difettoso dello stesso è una causa probabile del danno, o entrambi gli elementi".

Dunque, ci si riferisce, innanzitutto, ad una mera contribuzione (*il prodotto ha contribuito a cagionare il danno*), quindi, si fa riferimento ad un mero apporto del prodotto all'interno del nesso eziologico. L'effetto è che il testo proposto non si riferisce esclusivamente ad un nesso eziologico totalmente determinato, ma anche ad una "mera contribuzione", e, successivamente, la soglia della prova scende sino alla "probabilità" (*è probabile che il prodotto fosse difettoso, oppure che il carattere difettoso dello stesso è una causa probabile del danno*).

Quindi, nei casi di eccessiva difficoltà tecnica nella prova del difetto e del nesso di causalità tra difetto e danno, il Legislatore europeo si potrebbe accontentare (qualora il testo proposto fosse approvato) di una mera prova di probabilità.

A monte, peraltro, proprio per superare le richiamate asimmetrie informative e agevolare la posizione dell'attore, si sancisce anche un onere di divulgazione degli elementi di prova: *"gli Stati membri provvedono affinché gli organi giurisdizionali nazionali, su richiesta di un danneggiato che chiede il risarcimento del danno causato da un prodotto difettoso, sulla base di fatti e prove sufficienti a sostenere la plausibilità della domanda di risarcimento, abbiano il potere di ordinare al convenuto di divulgare i pertinenti elementi di prova a sua disposizione"*. Trattasi di un potere che deve essere esercitato nel rispetto del principio di proporzionalità onde tutelare eventuali segreti commerciali o industriali e, in generale, la riservatezza palesata dal produttore.

La proposta di direttiva punta ad alleviare l'onere della prova in capo al danneggiato e, nei casi complessi, ad accontentarsi di una sorta di probabilità e di un mero contributo causale, il tutto insieme all'introduzione di oneri di *disclosure* (cioè di fornire informazioni) a carico del produttore.

6.2. (Segue) *La proposta di direttiva sulla responsabilità per danni derivanti dall'output dell'IA: l'integrazione dei sistemi nazionali fondati sulla colpa*

L'intervento comunitario non si chiude qui in quanto a livello europeo si sta anche sviluppando un negoziato su un testo normativo specifico sulla responsabilità civile per danni dall'*output* prodotto dall'intelligenza artificiale.

Da questo punto di vista la proposta europea è assolutamente innovativa e, come le altre, rispetta, secondo la Commissione europea, i principi di proporzionalità e di sussidiarietà.

Il legislatore europeo punta a raggiungere un livello minimo di tutela idoneo ad integrare i sistemi nazionali.

Mentre prima si è ragionato con riguardo alla Direttiva prodotti e, dunque, in relazione alla responsabilità del produttore, che è un modello, secondo il diritto europeo, *fondato su una responsabilità oggettiva che concorre con la responsabilità contrattuale ricollegata alla vendita di beni di consumo, alla vendita di servizi e prodotti digitali, alla responsabilità in materia ambientale, alla responsabilità da illecito trattamento dei dati personali e concorre con i sistemi nazionali che non siano fondati su criteri di imputazione oggettiva*, in tal caso il legislatore comunitario si muove in un'altra area che si potrebbe definire *complementare*.

Il legislatore, infatti, si pone *nell'area della responsabilità fondata su un criterio di imputazione colposa* perseguendo l'obiettivo di integrare i sistemi nazionali attraverso una regolazione che punti soltanto a disciplinare due aspetti: 1) anche qui si punta all'eliminazione delle asimmetrie informative consentendo al danneggiato di ottenere elementi informativi idonei a perorare una causa risarcitoria; 2) inoltre si intende alleviare l'onere della prova. Il legislatore comunitario persegue questi due obiettivi affermando chiaramente che lo scopo della Direttiva è *migliorare il funzionamento del mercato interno stabilendo requisiti uniformi per determinati aspetti della responsabilità civile extracontrattuale per danni causati con il coinvolgimento di sistemi di intelligenza artificiale. La scelta di strumenti giuridici adeguati è limitata, date la natura della questione relativa all'onere della prova e le caratteristiche specifiche dell'intelligenza artificiale, che pongono problemi alle norme vigenti in materia di responsabilità. A tale riguardo la Direttiva alleggerisce della prova in modo molto mirato e proporzionato, attraverso il ricorso alla divulgazione e a presunzioni relative. In particolare, la Direttiva interviene in relazione alle domande di risarcimento del danno causate da un sistema di intelligenza artificiale, nel quadro di azioni civili di responsabilità extracontrattuale, qualora tali azioni siano intentate nell'ambito di regimi di responsabilità per colpa, ossia, in particolare, regimi che prevedono la responsabilità legale di risarcire i danni causati da un'azione o da un'omissione intenzionalmente lesiva o colposa. Le misure previste dalla Direttiva possono integrarsi senza attriti nei sistemi di responsabilità civile esistenti, in quanto riflettono un approccio che non incide sulla definizione di concetti*

fondamentali quali colpa o danno, dato che il significato di tali concetti varia notevolmente da uno Stato membro all'altro. Pertanto, al di là delle presunzioni da essa stabilite, la Direttiva lascia impregiudicate le norme dell'Unione o nazionali che stabiliscono, ad esempio, a quale parte incombe l'onere della prova e qual è il grado di certezza richiesto in relazione al livello della prova o che definiscono il concetto di colpa.

La proposta di direttiva, quindi, intende introdurre delle presunzioni per alleviare l'onere della prova, presunzioni che andranno ad operare all'interno degli ordinamenti nazionali senza cambiare le regole nazionali in materia di colpa o danno: *per quanto riguarda i danni causati dai sistemi di IA, la Direttiva mira a fornire una base efficace per le domande di risarcimento in relazione alla colpa, consistente nella conformità a un obbligo di diligenza a norma del diritto nazionale.*

È di estremo interesse richiamare le principali previsioni normative.

La presente Direttiva stabilisce norme comuni per quanto riguarda:

- la divulgazione di elementi di prova relativi a sistemi di intelligenza artificiale ad alto rischio, per consentire all'autore, in un'azione civile di responsabilità extracontrattuale per colpa, di motivare adeguatamente la domanda di risarcimento del danno;

- l'onere della prova nel quadro delle azioni civili di responsabilità extracontrattuale per colpa avviate dinanzi agli organi giurisdizionali nazionali, per ottenere il risarcimento del danno causato da un sistema di intelligenza artificiale.

La presente Direttiva si applica alle azioni civili di responsabilità extracontrattuale per colpa volte ad ottenere il risarcimento del danno causato da un sistema di intelligenza artificiale.

Non si applica, invece, alla responsabilità penale

Si tratta di una Direttiva di armonizzazione minima, in quanto gli Stati potranno introdurre disposizioni diverse e più protettive *(Gli Stati possono adottare o mantenere in vigore norme nazionali più favorevoli all'attore, in relazione all'esigenza di motivare una domanda presentata nel quadro di un'azione civile di responsabilità extracontrattuale, per ottenere il risarcimento del danno causato da un sistema di intelligenza artificiale, a condizione che tali norme siano compatibili con il diritto dell'Unione europea).*

Dunque, l'azione in relazione alla quale opereranno queste norme è quella risarcitoria, ossia una domanda proposta nel quadro di un'azione civile di responsabilità extracontrattuale per colpa, *volta ad ottenere il risarcimento del danno causato dall'output di un sistema di intelligenza artificiale, o dalla mancata produzione di un output che avrebbe, invece, dovuto essere prodotto da tale sistema.*

Quindi è un'azione di responsabilità per colpa, per danni derivanti dall'output prodotto dall'intelligenza artificiale o da un output omesso.

Anche in tal caso il legislatore prevede la c.d. *divulgazione degli elementi di prova.*

L'asimmetria informativa in danno del danneggiato viene attenuata consentendo agli Stati di imporre al produttore obblighi di *disclosure*, ossia di informazione, utili a sostegno della fondatezza dell'azione risarcitoria, ma senza modificare il regime dell'onere della prova (poiché l'attore dovrà comunque fornirla) e purché si fornisca un "principio di prova" (elemento richiesto per ottenere l'ordinanza di rivelazione). Infatti, si afferma che *gli Stati membri provvedono affinché, su richiesta di un attore potenziale che, dopo aver chiesto al fornitore, ad una persona soggetta agli obblighi del fornitore o ad un utente di divulgare gli elementi di prova pertinenti di cui dispone in relazione ad un determinato sistema di intelligenza artificiale ad alto rischio che si sospetta abbia cagionato il danno, si sia visto opporre un rifiuto, oppure su richiesta dell'attore, gli Organi giurisdizionali abbiano il potere di ordinare la divulgazione di detti elementi di prova da parte di tali persone. L'attore potenziale deve presentare, a sostegno di tale richiesta, fatti e prove sufficienti a sostenere la plausibilità della domanda di risarcimento del danno. Nel quadro di una domanda di risarcimento del danno gli Organi giurisdizionali ordinano ad una delle persone elencate al paragrafo 1 di divulgare gli elementi di prova solo se l'attore ha previamente compiuto ogni sforzo proporzionato per ottenere tali elementi di prova dal convenuto.*

Dunque, è previsto un ausilio per smussare l'asimmetria informativa, ma comunque si presuppone una condotta diligente da parte del potenziale attore.

Per ciò che riguarda la regolazione "a monte", dunque, si cerca di risolvere i problemi di asimmetria informativa consentendo l'esperimento delle azioni risarcitorie.

A valle, invece, quindi con riguardo all'azione risarcitoria esperita, la proposta di direttiva punta ad una attenuazione dell'onere della prova, introducendo una presunzione relativa del nesso di causalità in caso di colpa.

Nella negoziazione di questa Proposta c'era stata una certa opposizione ad introdurre presunzioni assolute e, allora, la Commissione ha deciso di ricorrere alle presunzioni relative, le quali consentono una prova contraria.

Si prevede che, *ai fini dell'applicazione delle norme in materia di responsabilità alle domande di risarcimento del danno, gli Organi giurisdizionali presumono l'esistenza del nesso di causalità tra la colpa del*

convenuto e l'output prodotto da un sistema di intelligenza artificiale, o la mancata produzione di un output da parte di tale sistema, se sono soddisfatte tutte le seguenti condizioni:

- *l'attore ha dimostrato, o l'Organo giurisdizionale ha presunto, la colpa del convenuto o di una persona della cui condotta il convenuto è responsabile, consistente nella non conformità ad un obbligo di diligenza previsto dal diritto nazionale e direttamente inteso a proteggere dal danno verificatosi;*
- *si può ritenere ragionevolmente probabile, sulla base delle circostanze del caso, che il comportamento colposo abbia influito sull'output prodotto dal sistema di intelligenza artificiale;*
- *l'attore ha dimostrato che il danno è causato dall'output prodotto dal sistema di intelligenza artificiale o dalla mancata produzione dell'output.*

Quindi, trattasi di una presunzione che scatta in presenza di una colpa, oppure di una dimostrazione che il danno sia stato prodotto dall'output.

Inoltre, si afferma che, nel caso di *domanda di risarcimento del danno riguardante, invece, il sistema di IA ad alto rischio*, l'Organo giurisdizionale non applica la presunzione di cui trattasi, se il convenuto dimostra che l'attore può ragionevolmente accedere ad elementi di prova e ha competenze sufficienti per dimostrare l'esistenza del nesso di causalità. In caso di *domanda di risarcimento del danno riguardante un sistema di intelligenza artificiale, che non è ad alto rischio*, la presunzione si applica solo se l'Organo giurisdizionale ritiene eccessivamente difficile per l'attore dimostrare l'esistenza del nesso di causalità.

Dalla proposta di direttiva specifica emerge l'idea di attenuare l'onere della prova, ma sempreché l'attore si sia, in qualche misura, attivato. Ove il convenuto dimostri, invece, l'esistenza della possibilità di fornire la prova, senza ricorrere alla *disclosure*, o in base alle competenze possedute dal danneggiato, in tal caso non opererà la presunzione prevista dalla proposta di direttiva.

6.3. (Segue) Il quadro normativo eurounitario nel complesso: conclusioni

Seguendo un approccio multilivello dalle politiche europee *de iure condito* e *de iure condendo* emerge che il Legislatore eurounitario sta muovendo su vari fronti.

Innanzitutto, con l'EU IA Act 2024, pur sostenendo e sollecitando lo sviluppo dell'IA in vari ambiti, è necessario considerare l'intelligenza artificiale un vero e proprio prodotto, da rendere sicuro a tutela dei diritti fondamentali della persona, evitando discriminazioni (algoritmi e IA con *output* discriminatori), nonché forme di controllo generalizzato di tipo biometrico al di fuori dei casi eccezionali in cui esso possa essere giustificato dalla necessità di tutelare più importanti interessi generali (tutela penale di alcuni beni). Si introduce, così, un sistema di controllo *ex ante* ed *ex post* sul prodotto "intelligenza artificiale".

Inoltre, l'UE si è interessata ai profili di responsabilità proponendo sia modificando la direttiva sulla responsabilità del produttore che introducendo norme specifiche sulla responsabilità per danni derivanti dall'output dell'IA.

Nella seconda area di intervento, quindi, facendo rientrare l'intelligenza artificiale nel concetto di prodotto, s'intende applicare la normativa (modificata ed integrata) sulla responsabilità del produttore ai danni prodotti dall'IA. Onde superare problemi di asimmetria informativa, le norme vengono anche rivisitate attenuando l'onere della prova nelle ipotesi a responsabilità complessa (tra cui figura quella nascente dall'IA) nonché introducendo ordini giudiziali di condivisione di informazioni utili all'esperimento dell'azione risarcitoria (sanzionando ogni omissione con l'operatività di una presunzione avente ad oggetto elementi costitutivi della responsabilità).

Nella terza area di intervento il Legislatore manifesta la volontà di integrare i sistemi nazionali di responsabilità fondati sulla colpa onde alleviare l'onere della prova, introducendo, a determinate condizioni, presunzioni e, sempre per risolvere i problemi di asimmetria informativa, prevedendo che i giudici nazionali possano imporre di rilasciare determinate informazioni a favore di attori futuri, potenziali ed eventuali.

Questa, in sintesi, è la strategia emergente dalle nuove linee di politica legislativa europea.

7. Il disegno di legge governativo italiano dell'aprile 2024: campo di applicazione, definizioni e principi generali. Amministrare e giudicare con algoritmi e IA?

Anche il Legislatore italiano è intervenuto in materia di IA con un disegno di legge governativo che si occupa di IA in modo trasversale introducendo norme in materia sanitaria, lavoristica, autorale, di giustizia ed amministrazione.

È un disegno che occupa spazi e aree che non verranno disciplinati dal diritto eurounitario approvato e approvando.

Fissa principi, concetti e definizioni.

Ad esempio, si precisa che l'IA è *un sistema automatizzato, progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output, quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali.*

I modelli di IA sono definiti quali *modelli che identificano strutture ricorrenti attraverso l'uso di collezione di dati, che hanno la capacità di svolgere un'ampia gamma di compiti distinti e che possono essere integrati in varietà di sistemi e applicazione.*

Come è bene evidente leggendo il testo proposto, il Legislatore segue un approccio molto simile a quello europeo: non intende ingessare lo sviluppo dell'intelligenza artificiale, ma guidarlo e regolarlo fissando alcuni principi generali.

Si prevede che *la ricerca, la sperimentazione, lo sviluppo, l'adozione, l'applicazione e l'utilizzo di sistemi di modelli di intelligenza artificiale avvengono nel rispetto dei diritti fondamentali e delle libertà previste dalla Costituzione, del diritto dell'Unione europea e dei principi di trasparenza, proporzionalità, sicurezza, valorizzazione, anche economica, del dato, protezione dei dati personali, eticità, riservatezza, robustezza, accuratezza, non discriminazione, parità di sessi e sostenibilità. Lo sviluppo dei sistemi e di modelli di intelligenza artificiale avviene sui dati e tramite processi, cui deve essere garantita e vigilata la correttezza, l'attendibilità, la sicurezza, la qualità, l'appropriatezza e la trasparenza, secondo i principi di proporzionalità e in relazione ai settori nei quali sono utilizzati. I sistemi e i modelli dell'intelligenza artificiale devono essere sviluppati ed applicati nel rispetto dell'autonomia e del potere decisionale dell'uomo, della prevenzione del danno, della conoscibilità, della spiegabilità e dei principi del primo comma.*

In più, *l'utilizzo di sistemi di intelligenza artificiale nell'informazione avviene senza pregiudizio alla libertà e al pluralismo dei mezzi di comunicazione, alla libertà di espressione, alla obiettività, completezza, imparzialità e lealtà della informazione. L'utilizzo di intelligenza artificiale garantisce il trattamento lecito, corretto, trasparente dei dati personali e la compatibilità con le finalità per le quali sono raccolti, in conformità con il diritto dell'Unione europea in materia di dati personali e di tutela della riservatezza.*

Sono menzionati anche altri principi, tra cui quelli di sviluppo economico, cybersicurezza e sicurezza nazionale.

Fondamentale è il principio del servizio: *l'intelligenza artificiale è vista come un quid a servizio dell'uomo. Questo principio emerge specialmente nelle norme in materia di lavoro, professioni intellettuali, di Pubblica amministrazione e attività giudiziaria. Si afferma: l'intelligenza artificiale è al servizio della persona ed è impiegata per migliorare le condizioni di lavoro, tutelare l'integrità psicofisica dei lavoratori, accrescere le qualità delle prestazioni lavorative e la produttività delle persone. Non può svolgersi in contrasto con la dignità umana, né violare la riservatezza dei dati personali. Inoltre, sono vietate le discriminazioni. In materia di professioni intellettuali, si precisa che l'utilizzo di sistemi di intelligenza artificiale, nelle professioni intellettuali, è consentito esclusivamente per esercitare attività strumentali e di supporto all'attività professionale richiesta e con prevalenza del lavoro intellettuale oggetto della prestazione d'opera.*

Sono enunciati anche principi in materia di organizzazione e azione amministrativa: *le pubbliche amministrazioni utilizzano l'intelligenza artificiale allo scopo di incrementare l'efficienza della propria attività, ridurre i tempi di definizione dei procedimenti e aumentare la qualità e la quantità dei servizi erogati ai cittadini e alle imprese, assicurando agli interessati la conoscibilità del suo funzionamento e la tracciabilità del suo utilizzo. L'utilizzo dell'intelligenza artificiale avviene in funzione strumentale e di supporto a quella umana, in ogni caso nel rispetto dell'autonomia e del potere decisionale della persona, che resta l'unica responsabile dei provvedimenti e dei procedimenti in cui sia stata utilizzata l'intelligenza artificiale.*

Nel caso di utilizzo dell'intelligenza artificiale nell'attività giudiziaria, è previsto che *i sistemi di intelligenza artificiale sono utilizzati per l'organizzazione e la semplificazione del lavoro giudiziario, per la ricerca giurisprudenziale e dottrinale, anche finalizzata all'individuazione di orientamenti interpretativi per la predisposizione di bozze di provvedimenti e per ogni altro impiego strumentale e di supporto all'attività giudiziaria.*

Anche dal punto di vista della sanità, ad esempio, l'intelligenza artificiale riveste una funzione di supporto e vi sono varie norme dedicate a questa materia.

L'approccio seguito dal Legislatore italiano è molto chiaro: l'intelligenza artificiale non è sostitutiva, ma si pone a servizio dell'essere umano. L'obiettivo è quello di migliorare la qualità del lavoro, l'efficienza del lavoro e della pubblica amministrazione (ci si muove nella prospettiva dell'amministrazione di risultato non discriminatoria e del buon andamento amministrativo, nel senso del rispetto della imparzialità amministrativa e della "buona amministrazione" per utilizzare le aggettivazioni costituzionali e eurounitarie). L'intelligenza artificiale potrà, ad esempio, snellire i processi decisionali (sono allo studio una serie di procedimenti per l'assistenza del magistrato). Grazie al Codice dell'amministrazione digitale e alle riforme processuali oggi è già tutto digitalizzato e per questo si pensa di "dare in pasto" all'IA il fascicolo telematico digitalizzato a fini di studio, ricerca e proposta decisionale (la funzione di ausilio dell'IA potrebbe essere svolta sia alla fase documentale che a quella di discussione orale, atteso che, attraverso l'IA è oramai possibile verbalizzare e trascrivere immediatamente ciò che viene detto oralmente). In tal modo, escludendosi la via del giudice-robot (che violerebbe il principio di "riserva di umanità"), l'IA dovrebbe limitarsi a svolgere attività di studio e ricerca, nonché, al più, formulare bozze di atti giudiziari, relazioni e decisioni. In tal modo l'IA svolgerebbe

esclusivamente attività strumentali di ausilio e supporto, fermo restando che la decisione finale dovrà essere della persona fisica cui verrà imputata.

Sul presupposto della rassicurante strumentalità dell'IA all'azione umana, il messaggio che lancia il Legislatore nel disegno di legge governativo è rispettoso dei principi di servizio e di riserva di umanità che impone che ogni decisione debba essere sempre assunta da una persona fisica, cosa che consente l'operatività dei controlli, anche giurisdizionali, e l'imputazione delle responsabilità connesse all'esercizio dei poteri amministrativo e giudiziario.

È noto che, dal punto di vista del diritto amministrativo, è sorto oramai da anni un vivo dibattito dottrinale e giurisprudenza sulla c.d. decisione algoritmica in cui si estrinseca l'esercizio *robotizzato* del potere. Dottrina e giurisprudenza hanno sdoganato l'utilizzo dell'algoritmo (c.d. decisioni automatizzate) nell'ambito dei procedimenti decisionali della pubblica amministrazione soprattutto nei procedimenti vincolati. C'è chi giunge persino ad auspicare l'estensione delle c.d. "decisione automatizzata algoritmiche" all'esercizio di poteri tecnico-discrezionali e a discrezionalità amministrativa pura (questa idea è oggi ancora minoritaria in dottrina e poco seguita in giurisprudenza).

Accanto a questa limitazione alla decisione *robotizzata* basata su tipologia e natura giuridica del potere amministrativo esercitato, emerge anche un'altra cautela introdotta dalla dottrina e sulla quale la giurisprudenza ha orientato la propria attenzione: la necessità che venga rispettato il *principio di trasparenza con piena accessibilità all'algoritmo*. Atteso che la decisione amministrativa deve essere motivata e sindacabile dal giudice, è essenziale che siano conoscibile l'algoritmo (ovverosia i criteri di programmazione dell'algoritmo e quelli su cui ha fondato la decisione).

La giurisprudenza civile ed amministrativa si è pronunciata con riguardo alle assegnazioni algoritmiche delle sedi di servizio agli insegnanti. Nell'ambito di questo contenzioso è emerso come l'Amministrazione avesse in realtà fatto assumere la decisione all'algoritmo, in modo automatizzato, ma creando esiti illogici, irragionevoli e discriminatori. Il problema si è altresì posto in relazione alla liquidazione di stipendi e pensioni, nonché in relazione all'esercizio di poteri fiscali dell'Amministrazione finanziaria. Oramai in tanti settori l'Amministrazione si affida all'algoritmo ossia all'intelligenza artificiale.

La giurisprudenza non ha manifestato un atteggiamento chiuso all'utilizzo delle nuove tecnologie. Le Corti hanno aperto all'utilizzo di algoritmi e IA, ma in una prospettiva molto simile a quella evidenziata dal Governo: deve sempre esistere -al termine del processo- una "decisione umana" e, in definitiva, la riferibilità dell'esercizio del potere ad un essere umano, anche affinché la macchina non venga utilizzata come mezzo, strumento o schermo per escludere ed eludere le responsabilità riconnesse alla *mala gestio* amministrativa.

Dunque, da una parte, è necessaria la riferibilità ad un essere umano (il processo decisionale deve essere controllato e supervisionato da un essere umano), dall'altra, deve sempre essere garantita l'accessibilità ai criteri su cui si fonda l'algoritmo e in virtù dei quali avvengono i processi decisionali algoritmici.

Il giudice deve poter sempre valutare e sindacare l'azione amministrativa, anche se svolta con il supporto dell'intelligenza artificiale: i criteri ideati dal programmatore del *software* per sviluppare l'algoritmo devono essere logici, ragionevoli, non discriminatori e sempre accessibili al destinatario dell'azione amministrativa e, in ultima istanza, al giudicante o all'organo di controllo (come noto, nella vicenda relativa all'assegnazione delle sedi agli insegnanti, l'algoritmo aveva dato luogo ad esiti del tutto abnormi, illogici, irragionevoli, ossia a risultati che si ponevano all'opposto del divisato obiettivo di indipendenza, efficienza e imparzialità). Ecco perché oggi il problema che pone l'utilizzo dell'intelligenza artificiale investe la "sua gestione" e cioè l'organizzazione di risorse umane capaci di controllare la macchina, supervisionare l'intelligenza artificiale, al fine di rendere decisioni amministrative ragionevoli, logiche, giuste e corrette, nonché sindacabili da parte del giudice.

Il disegno di legge governativo sul presupposto di queste acquisizioni dottrinali e giurisprudenziali intervenute con riguardo alle decisioni algoritmiche prevede che *le pubbliche amministrazioni utilizzano l'intelligenza artificiale allo scopo di incrementare l'efficienza della propria attività, di ridurre i tempi di definizione dei procedimenti e di aumentare la qualità e la quantità dei servizi erogati ai cittadini e alle imprese, assicurando agli interessati la conoscibilità del suo funzionamento e la tracciabilità del suo utilizzo*. Si sposa una prospettiva servente, con garanzia di accesso all'algoritmo e relativi sistemi di funzionamento, nonché di tracciabilità dell'utilizzo. Al cittadino o all'impresa deve essere noto che la decisione sia stata assunta attraverso meccanismi di intelligenza artificiale.

Si sancisce altresì che *l'utilizzo della intelligenza artificiale avviene in funzione strumentale e di supporto a quella umana. In ogni caso, nel rispetto dell'autonomia e del potere decisionale della persona, che resta l'unica responsabile dei provvedimenti e dei procedimenti in cui sia stata utilizzata l'intelligenza artificiale*. Emerge così il tema delle responsabilità, essendo necessari autonomia e potere decisionale della persona fisica, unica responsabile anche nei procedimenti in cui sia stata utilizzata l'intelligenza artificiale.

A tal proposito il disegno di legge governativo si ricongiunge con le più recenti previsioni di legge in materia di contratti pubblici e azione di accertamento dell'A.F. L'art. 30 d.lgs. 31 marzo 20023, n. 36, difatti,

consente alle stazioni appaltanti l'utilizzo dell'IA nel rispetto dei principi di trasparenza (conoscibilità, comprensibilità e accessibilità dell'algoritmo, del codice sorgente e dei criteri di funzionamento), di non esclusività della decisione elettronica automatizzata (principio di riserva di umanità che postula l'esistenza di un contributo umano capace di controllare, validare ovvero smentire la decisione automatizzata) e di non discriminazione.

Anche in materia di riforma fiscale (art. 17, c. 1, lett. f, d.lgs. 9 agosto 2023, n. 111) tra i criteri di delega è presente quello che consente all'A.F. di potenziare l'utilizzo di tecnologie digitali, anche con l'impiego di sistemi di intelligenza artificiale. Dal punto di vista dell'esercizio dei poteri amministrativi, dunque, il Legislatore asseconda il *sentiment* volto all'apertura all'intelligenza artificiale (sempre in una prospettiva di servizio).

Con riguardo all'attività giudiziaria, invece, il legislatore definisce e circoscrive con cautela gli ambiti di utilizzo dell'intelligenza artificiale, chiarendo che *i sistemi di intelligenza artificiale sono utilizzati per l'organizzazione e la semplificazione del lavoro giudiziario*. Dunque, l'utilizzo dei sistemi di intelligenza artificiale è normato in una prospettiva di organizzazione e di semplificazione, nonché *per la ricerca giurisprudenziale e dottrinale, anche finalizzata alla individuazione di orientamenti interpretativi* (che è quello che già fanno i Magistrati), *per la predisposizione di bozze di provvedimenti e per ogni altro impiego strumentale e di supporto alla attività giudiziaria*. Emerge una prospettiva di supporto e di strumentalità, ma il Legislatore si riferisce anche alla *predisposizione di bozze di provvedimenti*, ossia proprio delle attività prima indicate di immissione del fascicolo digitalizzato nei sistemi di IA persino per ottenere bozze di atti. Il disegno di legge chiarisce altresì che il Ministero della Giustizia debba regolare *l'impiego dei sistemi di intelligenza artificiale da parte degli Uffici giudiziari*, precisando che *è sempre riservata al Magistrato la decisione sulla interpretazione della legge, sulla valutazione dei fatti e delle prove e sulla adozione di ogni provvedimento*. Tale previsione è chiaramente rivolta alla tutela dell'interesse della collettività dei cittadini e delle imprese atteso che l'attività magistratuale è un servizio che il Magistrato svolge per gli altri alla stregua di una missione. Autorevole dottrina, seppure molto aperta ed interessata all'intelligenza artificiale, ha acutamente puntato l'attenzione verso il massimo rispetto del principio di servizio e riserva di umanità specie in relazione all'attività giudiziaria, di guisa che l'azione dell'IA non sia mai sostitutiva dell'essere umano.

7.1. (Segue) IA e tutela della proprietà intellettuale

Se si confronta il disegno di legge governativo italiano con l'*executive order* di Biden del 2023 emerge una sostanziale differenza: il Presidente degli Stati Uniti, in quell'atto, diede indicazioni operative a tutti i Dipartimenti affinché identificassero, in tutti i settori dell'amministrazione federale, come sviluppare e disciplinare l'intelligenza artificiale. Nel disegno di legge italiano, invece, si palesa una visione più limitata e circoscritta, con un interesse limitato ad alcuni settori specifici.

Il legislatore si occupa, però, anche di tutela del diritto di autore.

La norma di riferimento è stata già richiamata: essa introduce un criterio di bilanciamento degli interessi in gioco anche attraverso la necessità di verificare "la parte" attribuibile all'azione della macchina (o IA) e quella imputabile al contributo della persona fisica che pretende di essere l'autore di quell'opera intellettuale originale e innovativa.

Il disegno opera anche in un altro senso: occupandosi del contrasto delle cc.dd. *fake news*, si stabilisce che il contenuto che venga veicolato in determinati contesti debba recare una espressa indicazione esplicitante che esso sia stato realizzato attraverso l'ausilio dell'intelligenza artificiale. L'art. 40-bis sancisce che *qualunque contenuto informativo diffuso da fornitori di servizi audiovisivi e radiofonici, tramite qualsiasi piattaforma, in qualsiasi modalità, incluso il video on demand e lo streaming che, sempre previa acquisizione del consenso del titolare dei diritti, sia stato generato, modificato, alterato, anche parzialmente, in qualsiasi forma o modo attraverso l'utilizzo di sistemi di intelligenza artificiale, in grado di rappresentare come reali dati, fatti, informazioni che non lo siano, deve essere reso, a cura dell'autore o del titolare dei diritti di sfruttamento economico, se diverso dall'autore, chiaramente visibile e riconoscibile da parte degli utenti, mediante l'inserimento di un elemento o segno identificativo, anche in filigrana, purché facilmente visibile, con l'acronimo AI, ovvero, nel caso audio, attraverso annunci audio, ovvero con tecnologie adatte a consentire il riconoscimento*. Tutto questo per dare un *alert*, ossia per far capire ai destinatari di quell'informazione che il prodotto audio o audiovisivo sia stato realizzato anche attraverso meccanismi di intelligenza artificiale.

7.2. (Segue) IA e diritto penale: aggravanti e il nuovo reato di illecita diffusione di contenuti generati o manipolati artificialmente

Fino a questo momento è stata esaminata l'IA con riguardo al diritto civile e amministrativo, svolgendo talune importanti considerazioni anche in relazione alla tutela della proprietà industriale e intellettuale. È giunto il momento di passare al diritto penale atteso che il disegno di legge governativo, oltre ad introdurre una circostanza aggravante comune, prevede più aggravanti speciali.

Per esempio, l'art. 61 viene modificato introducendo un c. 11-*decies*, il quale prevede la circostanza aggravante dell'*aver commesso il fatto mediante impiego di sistemi di intelligenza artificiale quando gli stessi, per la loro natura e per le modalità di utilizzo, abbiano costituito mezzo insidioso, ovvero quando il loro impiego abbia, comunque, ostacolato la pubblica o privata difesa, ovvero aggravato le conseguenze del reato*.

Inoltre, il Legislatore introduce altre aggravanti in relazione a diverse fattispecie di reato. Ad esempio, in relazione alla *sostituzione di persona*, il reato viene aggravato *se il fatto sia commesso mediante l'impiego di sistemi di intelligenza artificiale*. Anche in relazione al *ribasso e al rialzo fraudolento di prezzi sul pubblico mercato nelle borse in commercio* viene previsto l'ipotesi aggravata del *fatto commesso mediante l'impiego di sistemi di intelligenza artificiale*. Nello stesso senso, si interviene in materia di truffa, di frode informatica, di riciclaggio, di reimpiego di denaro o di altre utilità di provenienza illecita, di autoriciclaggio ed aggio e, addirittura, in relazione al reato di *insider trading*.

Il legislatore migliora il sistema penale, tenendo in considerazione le condotte poste in essere attraverso l'utilizzo della intelligenza artificiale, introducendo delle aggravanti oppure consentendo l'integrazione del reato anche qualora la condotta venga posta in essere l'utilizzo della intelligenza artificiale.

In più, il legislatore propone l'introduzione, nell'ambito del Codice penale, di una nuova fattispecie di reato, ossa dell'art. 612-*quater*. Si è all'interno dei delitti a tutela della libertà morale. Tale reato, difatti, è posto dopo la *violenza privata, la minaccia, gli atti persecutori, la diffusione di immagini e video a contenuto sessualmente esplicito*.

Dopo tali norme, quindi, il legislatore propone l'introduzione di una nuova fattispecie di reato, ossia l'art. 612-*quater*, recante *illecita diffusione di contenuti generati o manipolati artificialmente*. Il testo della norma prevede che *chiunque cagiona ad altri un danno ingiusto mediante invio, consegna, cessione, pubblicazione o, comunque, diffusione di immagini o video di persone o cose, ovvero voci o suoni, in tutto o in parte, falsi, generati o manipolati mediante l'impiego di sistemi di intelligenza artificiale, atti ad indurre in inganno sulla loro genuinità o provenienza, è punito con la reclusione da 1 a 5 anni. Il delitto è punibile a querela della persona offesa. Si procede, tuttavia, d'ufficio se il fatto è commesso con un altro delitto per il quale si deve procedere d'ufficio, ovvero se è commesso nei confronti di persone incapaci, per età o infermità, o di una pubblica autorità a causa delle funzioni esercitate*.

La fattispecie, innanzitutto, è ben descritta e delineata dal Legislatore, il quale modella e descrive puntualmente il fatto facendo riferimento ad un ventaglio applicativo di condotte. Pertanto, il legislatore colma possibili vuoti di tutela.

La condotta deve puntare al risultato identificato dal Legislatore (tale caratteristica è comune a tutte le condotte alternativamente identificate dal Legislatore) ovverosia nell'evento di danno previsto dalla stessa fattispecie incriminatrice (si dice *mediante invio, consegna, cessione, pubblicazione o, comunque, diffusione di immagini o video di persone o cose [...], atti ad indurre in inganno sulla loro genuinità o provenienza*).

Si tratta di un reato comune (si punisce *chiunque cagiona ad altri un danno ingiusto*) e di una fattispecie causalmente orientata, in quanto il legislatore punta a punire quelle condotte che, comunque, abbiano determinato un danno ingiusto (non qualsivoglia condotta, ma solo quelle condotte indicate dal legislatore che abbiano cagionato l'evento dannoso previsto).

È un reato a forma vincolata, in quanto il *cagionare il danno ingiusto* è penalmente rilevante qualora ci sia *l'invio, la consegna, la cessione, la pubblicazione o, comunque, la diffusione di immagini o video di persone e di cose, ovvero di voci o suoni, in tutto o in parte, falsi* (questo è un altro elemento centrale della fattispecie) *generati o manipolati mediante l'impiego di sistemi di intelligenza artificiale, atti ad indurre in inganno sulla loro genuinità o provenienza*.

Dunque, è necessaria, ai fini della punibilità, una falsità (del suono, del contenuto, dell'audio etc.), una falsità che derivi dall'utilizzo della intelligenza artificiale e una idoneità di questi atti ad indurre in inganno sulla genuinità o provenienza.

Dal punto di vista strutturale, il Legislatore utilizza una tecnica casistica, che evita lacune, tenendo sempre presente che il comportamento umano penalmente rilevante deve tendere ad una modificazione della realtà. Tale modificazione della realtà è *il danno ingiusto*, un danno ingiusto che, però, deve essere arrecato con queste particolari modalità.

Quindi, si tratta di una falsità generata dall'intelligenza artificiale, ma non deve essere una falsità macroscopica, grossolana, bensì una falsità che renda possibile l'inganno, ossia, come si esprime la norma, *idonea ad indurre in inganno sulla loro genuinità e provenienza*.

Ci si può chiedere quali siano i beni giuridici protetti. Probabilmente potrebbe trattarsi di un reato pluri-offensivo. Si deve anche decodificare il concetto di *danno ingiusto*. In una logica autonomistica e, dunque, non pancivile, il concetto di *danno ingiusto*, anche se rievoca l'art. 2043 c.c. (non l'art. 2059 c.c.), dovrebbe essere interpretato in senso lato, ossia in una prospettiva autonomistica.

Dovrebbe trattarsi di un danno riferibile non soltanto a pregiudizi patrimoniali, ma anche a quelli non patrimoniali. Il danno ingiusto dovrebbe essere quindi presente al cospetto una *lesione di situazioni soggettive giuridicamente rilevanti per l'Ordinamento penale* (nella specie per il nuovo art. 612-*quater*): accanto alla libertà morale, si potrebbe ritenere che il Legislatore tuteli tutte quelle situazioni soggettive incise dalla condotta, che potrebbero avere una rilevanza patrimoniale e non patrimoniale.

Potrebbe probabilmente emergere anche un altro bene giuridico protetto. L'esistenza del falso e il fatto che tutte queste condotte devono indurre in inganno sulla genuinità e sulla provenienza fanno sì che emerga un discorso tipicamente correlato alla fede pubblica e, quindi, a quell'affidamento della collettività nella genuinità e nella originalità di quella informazione. Non verrebbe tutelato, quindi, soltanto l'interesse della persona offesa, ma si tutelerebbe anche un interesse più ampio della collettività.

Un altro aspetto problematico è ricollegato all'inciso *generati o manipolati mediante l'impiego di sistemi di intelligenza artificiale*. In sostanza, immagini, video, suoni, voci, devono essere, in tutto o in parte, falsi, ma la falsità deve essere una falsità determinata dalla attività della intelligenza artificiale.

La falsità emerge dal fatto che quei suoni, quelle immagini e quei contenuti siano stati generati o manipolati dalla intelligenza artificiale. Questa è una ulteriore specificazione della condotta tipica penalmente rilevante in questo reato comune di evento causalmente orientato, dove il *danno ingiusto* deve essere determinato da questa condotta specifica.

Ciò che il legislatore non precisa – e sembra che lo faccia volutamente per estendere il penalmente rilevante – è che l'attività manipolativa debba essere stata effettuata da chi abbia cagionato il danno mediante la diffusione o mediante l'invio di questi contenuti. Il legislatore, in altri termini, non costruisce il fatto tipico imponendo che il soggetto agente, dopo aver utilizzato l'intelligenza artificiale, *diffonda*. In sostanza, si presuppone un intervento della intelligenza artificiale e da parte del soggetto attivo, o di un terzo, e l'importante è che vi sia, oggettivamente, un falso determinato dall'intelligenza artificiale.

Quindi, in questo modo, ciò che viene penalmente incriminato è proprio il fatto della diffusione e, leggendo l'altro segmento della norma (*atti ad indurre in inganno sulla loro genuinità e provenienza*), si comprende che l'invio, la consegna, la cessione, la pubblicazione o, comunque, la diffusione debbano essere condotte offensive (e, quindi, vi deve essere una offensività della condotta tipica).

Il fatto deve essere, cioè, un fatto penalmente rilevante, anche perché il reato non è introdotto (dal punto di vista della sistematica codicistica) nei *delitti contro la fede pubblica*. Esso è posto, a livello sistematico, nei *delitti contro la persona*. Quindi, il danno ingiusto non può che essere riferito ad una persona che venga ritratta, o la cui voce venga utilizzata, o la cui immagine e i cui tratti somatici siano usati generando video, suoni o immagini, oppure manipolando video, suoni o immagini (pertinenti a quella specifica persona).

Quindi, la condotta deve essere una condotta offensiva e il legislatore dipinge il fatto tipico nel pieno rispetto del principio di offensività, strutturando sicuramente un reato causalmente orientato, che punta ad un *danno ingiusto* (concetto che va decodificato, probabilmente in termini penalistici e, dunque, in termini autonomi rispetto al diritto civile), ma a forma vincolata: la condotta tenuta dall'agente deve assolutamente essere ad immagine e somiglianza di quella tipica descritta dal Legislatore, dove il presupposto della condotta tipica è quell'attività di falso, che rientra nella struttura del reato, e in relazione alla quale si porranno, come è possibile immaginare, tutti i problemi relativi ai rapporti con i delitti contro la fede pubblica.

Altro aspetto importante è quello che investe l'elemento soggettivo. È chiaro che quest'ultimo debba coprire tutti gli elementi tipici del fatto che si sono sino a questo momento descritti. Ma in tal caso vi è il solito problema da affrontare.

Infatti, nella parte in cui il Legislatore precisa che *tutte queste attività relative alla diffusione di immagini, video, voci, suono, in tutti o in parte falsi*, debbano essere *atte ad indurre in inganno sulla loro genuinità e provenienza*, significa che la condotta tipica deve essere atta a trarre in inganno e l'inganno deve avere ad oggetto la genuinità e la provenienza.

Gli esempi ben conosciuti sono video di un certo tipo, immagini che ritraggono cantanti, politici, personaggi famosi (addirittura uno *youtuber* si chiedeva pubblicamente se l'intelligenza artificiale potesse creare autonomamente video utilizzando i video propri e genuini presenti sul web, dando poi una risposta positiva al quesito richiamando un esperimento effettuato in un'università italiana).

Quindi, la condotta deve essere *atta ad indurre in inganno*, cioè, deve avere oggettivamente quei tratti idonei ad indurre in inganno. Sembra, quindi, che la norma, tra la “finalità dell'azione” e la “finalità dell'agente”, sterzi verso la prima. Cioè, non è tanto rilevante il fine che muove l'agente (certo, qualcuno potrebbe sostenere di essere al cospetto di una norma a dolo specifico con funzione selettiva del penalmente rilevante, tesi per la quale potrebbe rilevare la finalità dell'agente).

In realtà, però, essendo un reato causalmente orientato a forma vincolata, quindi, orientato al danno ingiusto, dove la condotta deve essere oggettivamente quella configurata dal legislatore, sembra che, nella prospettiva della offensività, questo segmento normativo, ossia l'*idoneità ad indurre in inganno*, debba essere interpretato come “finalità dell'azione”, ossia l'azione, oggettivamente, per le proprie caratteristiche,

per le proprie connotazioni e per come si manifesta nella fattispecie concreta, deve essere atta ad indurre in inganno e, dunque, dal punto di vista dell'elemento soggettivo autorevole dottrina è schierata per il dolo generico classico, senza che sia ipotizzabile alcun dolo specifico.

* * *

L'INTELLIGENZA ARTIFICIALE NEL DIRITTO AMMINISTRATIVO – TRA RISERVA DI UMANITÀ E NECESSITÀ DI GARANTIRE UNA MAGGIORE EFFICIENZA AMMINISTRATIVA

di Antonello Liroi (*)

Abstract: Lo sviluppo dell'Intelligenza artificiale in Europa e negli Stati membri passa inevitabilmente dall'integrazione dell'Intelligenza artificiale nel diritto amministrativo, esplorando il delicato equilibrio tra l'efficienza amministrativa e la salvaguardia della riserva di umanità. L'uso crescente dell'IA nei processi decisionali amministrativi promette di migliorare l'efficienza e la tempestività delle decisioni, riducendo i costi e il carico di lavoro del personale. Tuttavia, ciò solleva preoccupazioni riguardo alla trasparenza, all'imparzialità e alla responsabilità delle decisioni automatizzate, che devono essere conformi ai principi legali e garantire una supervisione umana efficace. La protezione dei dati personali e la sicurezza informatica emergono come ulteriori sfide da affrontare, insieme alla necessità di sviluppare linee guida chiare e meccanismi di controllo.

The development of artificial intelligence in Europe and its Member States inevitably involves integrating AI into administrative law, exploring the delicate balance between administrative efficiency and safe-guarding human dignity. The increasing use of AI in administrative decision-making promises to enhance the efficiency and timeliness of decisions, reducing costs and workload for personnel. However, this raises concerns regarding transparency, impartiality, and accountability of automated decisions, which must adhere to legal principles and ensure effective human oversight. Additional challenges include protecting personal data and ensuring cybersecurity, necessitating the development of clear guidelines and control mechanisms. Addressing these challenges effectively will require collaboration among legal professionals, AI experts, and policymakers to balance technological innovation with the protection of fundamental rights and administrative justice.

Sommario: 1. Cosa si intende per Intelligenza artificiale e quali benefici può portare. – 2. La trasparenza delle decisioni adottate da sistemi di intelligenza artificiale e i principali problemi di questo strumento. – 3. La risposta Europea ai rischi intrinseci all'uso dell'IA. – 4. Il ddl governativo italiano dell'aprile 2024: approccio e spunti principali. – 5. Il rapporto tra intelligenza artificiale con la disciplina penalistica e la tutela della proprietà intellettuale. – 6. Il concetto di pubblica amministrazione come “Casa di Vetro”. – 7. Il quadro normativo italiano in materia automazione e digitalizzazione. – 8. L'intelligenza artificiale e le nuove soluzioni tecnologiche applicate al ciclo di vita dei contratti pubblici. 9. Bilanciamento dei principi sottesi all'agire amministrativo. Una nuova interpretazione costituzionalmente orientata. – 10. Possibili evoluzioni future delle nuove tecnologie nel settore dei contratti pubblici. – 11. Considerazioni conclusive.

1. Cosa si intende per Intelligenza artificiale e quali benefici può portare

“L'intelligenza artificiale (IA) è una disciplina dell'informatica che sviluppa sistemi capaci di svolgere compiti che normalmente richiedono intelligenza umana, come il riconoscimento di immagini, la comprensione del linguaggio naturale, e la risoluzione di problemi complessi. Utilizza tecniche come l'apprendimento automatico, le reti neurali artificiali e l'elaborazione del linguaggio naturale per migliorare continuamente le proprie prestazioni attraverso l'esperienza e i dati. L'IA è applicata in vari settori, migliorando l'efficienza e creando nuove opportunità in campi come la sanità, la finanza e l'automazione industriale”.

La definizione di cui sopra, a prescindere dalla sua bontà o completezza, non rappresenta il pensiero dell'autore o della dottrina. O meglio, potrebbe essere la sintesi di tutte le definizioni e valutazioni dell'Intelligenza artificiale reperibili tramite ricerca *internet*.

La definizione sopra, infatti, è stata creata con un sistema di Intelligenza artificiale a cui ci si è limitati a chiedere “descrivi in cinque righe – e nel modo più completo possibile – l'Intelligenza artificiale”.

Sebbene tale definizione possa sembrare completa, non sono in grado di comprendere o visionare i ragionamenti che il sistema ha fatto per dare quella risposta, che senza dubbio sarà guidata dalle informazioni e stimoli che il sistema, nel caso di specie la popolare piattaforma “ChatGpt” ha ricevuto sin dalla sua programmazione.

Una definizione di Intelligenza artificiale, però, viene fornita – dal punto di vista normativo – dal nuovo regolamento europeo in materia di Intelligenza artificiale (di cui si parlerà *infra*) il quale sancisce che: “con il termine *intelligenza artificiale* si indica una famiglia di tecnologie, in rapida evoluzione, in grado di apportare una vasta gamma di benefici economici e sociali in tutto lo spettro delle attività industriali e sociali. L'uso dell'intelligenza

(*) A. Liroi, è Avvocato, partner dello Studio Gianni & Origoni: Co-Responsabile del dipartimento di Diritto amministrativo, che include anche i settori ambientale, *public utilities*, società pubbliche.

artificiale, garantendo un miglioramento delle previsioni, l'ottimizzazione delle operazioni e dell'assegnazione delle risorse e la personalizzazione dell'erogazione dei servizi, può contribuire al conseguimento di risultati vantaggiosi dal punto di vista sociale e ambientale, nonché fornire vantaggi competitivi fondamentali alle imprese e all'economia europea" (1).

Dalla definizione di cui sopra emerge, *inter alia*, il valore dell'Intelligenza artificiale quale strumento idoneo a contribuire ad importanti risultati in svariati ambiti della vita quotidiana dei cittadini, oltre a poter intervenire in moltissimi settori di grande importanza economica e sociale.

Per avere un'idea di quali sono le potenzialità dei sistemi di Intelligenza artificiale, al di fuori della sfera giuridica su cui si discuterà successivamente nel testo, si pensi che in ambito sanitario, l'IA supporta la diagnosi precoce di malattie e personalizza i piani di trattamento, migliorando così l'efficacia delle cure; nella mobilità urbana, algoritmi avanzati ottimizzano il traffico e migliorano i servizi di trasporto pubblico, riducendo i tempi di viaggio e l'impatto ambientale; nel settore finanziario, l'IA automatizza la gestione delle finanze personali e la rilevazione di frodi, offrendo maggiore sicurezza e semplificazione per gli utenti e, inoltre; in ambito educativo, le piattaforme di apprendimento personalizzato basate su IA adattano i contenuti alle esigenze specifiche degli studenti, migliorando l'efficacia dell'insegnamento (2).

In questa sede preme, però, analizzare l'impatto dell'uso dell'Intelligenza artificiale unicamente con riferimento all'agire amministrativo. Difatti, la pubblica amministrazione sta già attraversando quella che è una sorta di quarta rivoluzione industriale che mira ad ottenere un elevato grado di interconnessione e automazione dei diversi procedimenti amministrativi parte dell'azione pubblica. Ciò pone, in ogni caso, la pubblica amministrazione di fronte a sfide particolarmente complesse, che vanno ben oltre il mero passaggio dalla gestione di file cartacei a file digitali.

Lo scopo di questo contributo è, difatti, analizzare in che modo i processi decisionali in ambito amministrativo, che sono spesso e volentieri guidati da forte discrezione da parte dell'Autorità competente, possono essere delegati efficacemente a sistemi di Intelligenza artificiale, senza che ciò pregiudichi i diritti dei cittadini, valutando come l'Intelligenza artificiale sia uno strumento particolarmente utile ma che porta con sé una serie di rischi intrinseci dovuti – *inter alia* – alla possibile assenza di trasparenza delle decisioni adottate dall'IA.

2. La trasparenza delle decisioni adottate da sistemi di intelligenza artificiale e i principali problemi di questo strumento

I benefici portati dall'Intelligenza artificiale portano, come ogni strumento rivoluzionario, molteplici rischi intrinseci connessi sia ai possibili usi *contra legem* che si possano fare (e di cui si discuterà *infra*) sia a caratteristiche proprie dello strumento.

In tal senso, il principale problema si crea in relazione a quei sistemi di Intelligenza artificiale – c.d. *generative IA* – il cui uso si presta soprattutto in relazione alle professioni intellettuali e il cui *output* si basa sia sugli *input* forniti dal gestore del *software*, ma anche su attività di programmazione che lo stesso sistema svolge automaticamente.

Il problema è particolarmente noto, tanto che lo stesso Consiglio dell'Unione europea esortava le istituzioni comunitarie “*ad affrontare l'opacità, la complessità, la faziosità, un certo grado di imprevedibilità e un comportamento parzialmente autonomo di taluni sistemi di IA, onde garantirne la compatibilità con i diritti fondamentali e agevolare l'applicazione delle norme giuridiche*”.

Tale opacità è conosciuta anche come effetto “*black box*” e concerne la difficoltà, o addirittura l'impossibilità, di comprendere il processo decisionale di un sistema di IA, a causa della complessità e opacità degli algoritmi utilizzati. Quando un'IA prende una decisione, spesso non è chiaro come e perché sia giunta a quella specifica conclusione, rendendo difficile per gli esseri umani interpretare, contestare o migliorare tali decisioni.

L'effetto *black box* è particolarmente problematico nel contesto giuridico, dove la trasparenza e l'*accountability* sono fondamentali per garantire la legittimità e la giustizia dei processi decisionali, ciò che potrebbe portare a decisioni arbitrarie o discriminatorie, minando la fiducia dei cittadini nelle istituzioni e nei processi basati sull'uso di Intelligenza artificiale (3).

Inoltre, l'opacità delle ragioni che portano un sistema di Intelligenza artificiale ad assumere una determinata decisione può comportare una riduzione del diritto alla difesa dell'individuo il quale per contestare la decisione a lui relativa avrà la necessità di comprendere il procedimento, il ragionamento e gli elementi fattuali che l'hanno

(1) Considerando 1, Regolamento 1689/2024 “Artificial Intelligence Act” (2024).

(2) V. Corneli, *Sovranità tecnologica: intelligenza artificiale e valori costituzionali*, in <www.forumcostituzionale.it>, 4 giugno 2023.

(3) Frank Pasquale *The Black Box Society: The Secret Algorithms That Control Money and Information* (2015), Harvard University Press.

generata. Per garantire un equo processo, infatti, è essenziale che le parti interessate possano accedere alle logiche e ai criteri sottostanti le decisioni automatizzate (4).

Sotto questo aspetto, emerge come, la *generative AI* pone problematiche di natura etica e sociale, che rendono necessario mantenere un margine di intervento umano tale da mitigare l'intervento dell'IA.

Per tale ragione è emersa, a livello globale, la necessità di salvaguardare la c.d. "riserva di umanità", ovvero la necessità di mantenere un controllo umano – anche di ultima istanza – sul lavoro di un sistema di Intelligenza artificiale (5). Tale necessità è stata poi postulata nel Regolamento in materia di Intelligenza artificiale, il quale ha affermato il principio per cui debbano essere fissate alcune limitazioni anche al fine di assicurare l'operatività delle responsabilità e dei controlli, nonché la tutela delle situazioni soggettive coinvolte dall'esercizio dei poteri amministrativi e giudiziari attraverso strumenti di IA.

Da ciò, nasce l'esigenza di governare, dunque normare, l'utilizzo di questi nuovi strumenti digitali i cui impieghi sono ancora inesplorati, sia pur nella consapevolezza che, forse, nemmeno uno stretto principio di legalità potrebbe bastare a legittimare un'azione amministrativa "intelligente" ed automatizzata.

Certo è, che detta disciplina non può prescindere dal regolare tanto gli esiti ultimi dei sistemi algoritmici, quanto gli *input* iniziali che orientano le soluzioni delle *machine learning*.

3. La risposta europea ai rischi intrinseci all'uso dell'IA.

Vista la crescente importanza strategica dell'uso dell'Intelligenza artificiale in diversi settori nevralgici dell'economia, l'Unione europea (e anche alcuni Stati membri come l'Italia) ha deciso di regolamentare detto strumento. In particolare, l'Unione europea, partendo dall'idea che l'IA debba essere sviluppata e utilizzata in modo sicuro, etico e rispettoso dei diritti fondamentali e dei valori ha deciso di adottare un Regolamento denominato "*Artificial Intelligence Act*" ("AI Act" o il "Regolamento").

In particolare, l'AI Act è una proposta di regolamento presentata dalla Commissione europea il 21 aprile 2021, che mira a creare un quadro normativo armonizzato e proporzionato per l'Intelligenza artificiale nell'Unione europea. Il Regolamento è stato pubblicato sulla Gazzetta ufficiale dell'Unione europea in data 12 luglio 2024 e sarà applicabile 24 mesi dopo la data di pubblicazione.

Il Regolamento parte proprio dalla consapevolezza che lo strumento porti con sé molteplici rischi, affermando che "*gli stessi elementi e le stesse tecnologie che alimentano i benefici socio-economici dell'IA possono, altresì, comportare nuovi rischi o conseguenze negative per le persone giuridiche e la società*". In aggiunta, si riconosce come "*l'interesse dell'Unione è quello di preservare la leadership tecnologica dell'Ue e assicurare che i cittadini europei possano beneficiare delle nuove tecnologie sviluppate operanti in conformità ai valori, ai diritti fondamentali e ai principi dell'Unione*".

Per permettere quindi di godere di tali nuove tecnologie nel rispetto dei diritti fondamentali degli individui, il Regolamento richiama il principio di proporzionalità, assumendo la necessità di "*oneri normativi*" soltanto laddove un sistema di IA possa comportare rischi alti per i diritti fondamentali e la sicurezza. Per altri sistemi di IA non ad alto rischio sono imposti "*soltanto obblighi di trasparenza molto limitati, ad esempio in termini di fornitura di informazioni per segnalare l'utilizzo di un sistema di IA nelle interazioni con esseri umani*".

Appare quindi necessario regolamentare lo strumento avendo però quale obiettivo la tutela dei diritti fondamentali, come affermato nello stesso AI Act il quale, attraverso la definizione dei requisiti per una IA affidabile e di obblighi proporzionati per tutti i partecipanti alla catena del valore, mira a migliorare e promuovere la protezione dei diritti tutelati dalla Carta (6).

A tal fine il Regolamento prevede una classificazione dei sistemi di IA in base al loro livello di rischio per la sicurezza e i diritti delle persone, e stabilisce una serie di requisiti e obblighi per i fornitori e gli utenti di tali sistemi.

In particolare, il Regolamento distingue tra quattro categorie di rischio: inaccettabile, alto, limitato e minimo. In relazione ai vari livelli di rischio, il Regolamento stabilisce che:

- i sistemi di IA che contraddicono i valori e i principi fondamentali dell'Ue, come il rispetto della dignità umana, della democrazia e dello stato di diritto presentano un livello di rischio inaccettabile e, pertanto, sono

(4) C. Coglianese, D. Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, in *The Georgetown Law Journal*, 2017, vol. 105.

(5) F. Sarzana, S. Ippolito, M. Nicotra, *Diritto della Blockchain, Intelligenza artificiale e Iot*, 1ª ed., 2018, Ipsoa.

(6) Quali ad esempio: la protezione della vita privata, dei dati personali, della dignità umana, della parità di genere; si afferma la necessità di prevenire un effetto dissuasivo sui diritti di libertà di espressione e di riunione; si esplicitano numerose previsioni relative a rischi di decisioni assistite dalle IA errate o distorte in settori critici quali l'istruzione, la formazione, l'occupazione e le attività di contrasto di reati, dovendosi, in tutti questi casi, implementarsi regole idonee a prevenire ed evitare i predetti rischi).

vietati o fortemente limitati (come nel caso di sistemi di sorveglianza basati su analisi biometriche in tempo reale normalmente utilizzati per ragioni di pubblica sicurezza) (7).

- I sistemi di IA che possono avere un impatto significativo sui diritti fondamentali o sulla sicurezza delle persone presentano un livello di rischio elevato, ciò che impone che siano soggetti a rigorosi obblighi e requisiti individuati dal Regolamento prima di poter essere immessi sul mercato o utilizzati (8).

- I sistemi di IA che invece possono influenzare i diritti o le volontà degli utenti, ma in misura minore rispetto ai sistemi ad alto rischio sono invece considerati a rischio limitato. Tali sistemi sono soggetti a requisiti di trasparenza volti a favorire la trasparenza, permettendo agli utenti di essere consapevoli del fatto che interagiscono con un sistema di IA e di comprenderne le caratteristiche e le limitazioni (9).

- Presentano infine rischio minimo o nullo i sistemi di IA che non hanno alcun impatto diretto sui diritti fondamentali o sulla sicurezza delle persone, e che offrono ampi margini di scelta e controllo agli utenti. Questi sistemi sono liberi da qualsiasi obbligo normativo, al fine di incoraggiare l'innovazione e l'esperimentazione (10).

Sulla base di questa classificazione dei livelli di rischio, il Regolamento mira a garantire la sicurezza e l'eticità dell'IA, proteggendo i diritti e gli interessi delle persone interessate: i) definendo requisiti e gli obblighi per i sistemi di IA ad alto rischio, che assicurino il rispetto dei diritti fondamentali, quali il diritto alla dignità, alla non discriminazione, alla *privacy*, alla protezione dei dati, alla libertà di espressione e di informazione, al processo equo e alla presunzione di innocenza; ii) introducendo meccanismi di supervisione umana, che permettano alle persone di monitorare, controllare e intervenire sui sistemi di IA, al fine di prevenire o correggere eventuali effetti negativi o dannosi sull'essere umano o sull'ambiente; iii) la predisposizione di meccanismi di divieto assoluto per i sistemi di IA che presentano un rischio inaccettabile; iv) istituzione di un sistema di *governance* efficace e armonizzato per l'IA, che coinvolga tutte le parti interessate, tra cui le autorità nazionali competenti, la Commissione europea, il comitato europeo per l'IA, gli organismi di certificazione, i fornitori e gli utenti di sistemi di IA, e la società civile.

Questo sistema prevede la cooperazione, il coordinamento, il monitoraggio, la revisione, l'informazione, la consulenza, l'ispezione, l'azione correttiva e sanzionatoria.

A margine, il progetto europeo mira alla promozione di una cultura dell'IA responsabile e consapevole, che incoraggi la partecipazione, il dialogo, la trasparenza, la responsabilità e l'educazione degli sviluppatori, degli utilizzatori e dei cittadini di sistemi di IA, al fine di rafforzare la fiducia, il rispetto e l'accettazione sociale dell'IA.

4. Il ddl governativo italiano dell'aprile 2024: approccio e spunti principali

Anche il legislatore italiano è intervenuto in materia di IA con un disegno di legge governativo.

Il ddl – all'art. 2 – definisce un sistema di Intelligenza artificiale come: “*un sistema automatizzato, progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'imput che riceve come generare output, quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali.*”

L'approccio seguito dal legislatore nazionale appare molto simile a quello comunitario, con il ddl che dovrà completare il quadro già delineato dal legislatore europeo.

In tal senso, il ddl, così come il Regolamento, perseguono finalità di interesse pubblico e hanno lo scopo di creare regole volte allo sviluppo dei sistemi di Intelligenza artificiale per agevolare i principali aspetti della vita quotidiana.

A tal fine, il ddl indica, già al Capo II gli ambiti di applicazione specifici dell'Intelligenza artificiale relativamente ai settori di maggior interesse pubblico, dedicando un articolo specifico per ogni settore: sanitario; tutela del lavoro; professioni intellettuali; attività giudiziaria e pubblica amministrazione.

Il presupposto della norma appare particolarmente chiaro dal dettato dell'art. 10 del ddl in materia di tutela del lavoro, che riconosce che “*l'intelligenza artificiale è impiegata per migliorare le condizioni di lavoro, tutelare l'integrità psicofisica dei lavoratori, accrescere le qualità delle prestazioni lavorative e la produttività delle persone.*” e inoltre che “*l'utilizzo dell'intelligenza artificiale in ambito lavorativo [...] non può svolgersi in contrasto con la dignità umana, né violare la riservatezza dei dati personali.*”

Dall'art. 12, invece, che disciplina il rapporto degli strumenti di Intelligenza artificiale con le professioni intellettuali emerge il ruolo di strumento ausiliario riconosciuto a detti sistemi. Difatti, si afferma che *l'utilizzo di*

(7) Sono proibiti, *inter alia*, i sistemi di IA capaci astrattamente di manipolare il comportamento umano eludendone le reali volontà, ma anche quei sistemi che consentono il c.d. “*scoring sociale*” da parte delle autorità pubbliche.

(8) Un esempio di tali sistemi è costituito da quelli utilizzati nella selezione del personale, per l'ammissione all'istruzione, per l'erogazione di servizi sociali essenziali, come la Sanità, per la sorveglianza biometrica a distanza (non in tempo reale), per le applicazioni giudiziarie e di polizia, o per la gestione della sicurezza critica delle infrastrutture.

(9) Rientrano in questa categoria i sistemi di IA utilizzati per generare o manipolare contenuti audiovisivi (come i *deepfake*), o per fornire suggerimenti personalizzati.

(10) Rientrano in questa categoria i sistemi di IA utilizzati per scopi ludici o per scopi puramente estetici.

sistemi di intelligenza artificiale, nelle professioni intellettuali, è consentito esclusivamente per esercitare attività strumentali e di supporto all'attività professionale richiesta e con prevalenza del lavoro intellettuale oggetto della prestazione d'opera.

Ciò viene evidenziato nuovamente dal dettato concernente l'organizzazione e azione amministrativa di cui all'art. 13, che indicano come *“le pubbliche amministrazioni utilizzano l'intelligenza artificiale allo scopo di incrementare l'efficienza della propria attività, ridurre i tempi di definizione dei procedimenti e aumentare la qualità e la quantità dei servizi erogati ai cittadini e alle imprese, assicurando agli interessati la conoscibilità del suo funzionamento e la tracciabilità del suo utilizzo. L'utilizzo dell'intelligenza artificiale avviene in funzione strumentale e di supporto a quella umana, in ogni caso nel rispetto dell'autonomia e del potere decisionale della persona, che resta l'unica responsabile dei provvedimenti e dei procedimenti in cui sia stata utilizzata l'intelligenza artificiale.*

L'analisi del dettato dell'art. 13, pertanto, porta a due importanti riflessioni: da un lato, si evidenzia, con l'ultimo capoverso, il ruolo dell'individuo che utilizza i sistemi di Intelligenza artificiale e che, dal punto di vista giuridico, resta l'unico soggetto responsabile del provvedimento adottato – anche – tramite Intelligenza artificiale; dall'altro, il legislatore nazionale intende rimuovere le perplessità in merito ai principi e ragionamenti che hanno guidato la decisione dell'Intelligenza artificiale, riconoscendo che debba essere assicurata agli interessati la conoscibilità del suo funzionamento e la tracciabilità del suo utilizzo.

Rimanendo in ambito amministrativo, sia la dottrina che la giurisprudenza (11) si sono già espressi in merito all'utilizzo di decisioni automatizzate nell'ambito dei procedimenti decisionali della pubblica amministrazione, con particolare riferimento ai procedimenti vincolati, arrivando, in alcuni casi a valutare di estendere l'uso di decisioni automatizzate anche con riferimento a poteri tecnico-discrezionali e a discrezionalità amministrativa pura (di questo si parlerà *infra*).

Per quanto concerne, infine, l'uso dell'Intelligenza artificiale nell'attività giudiziaria, è previsto che *i sistemi di intelligenza artificiale sono utilizzati per l'organizzazione e la semplificazione del lavoro giudiziario, per la ricerca giurisprudenziale e dottrinale, anche finalizzata all'individuazione di orientamenti interpretativi per la predisposizione di bozze di provvedimenti e per ogni altro impiego strumentale e di supporto all'attività giudiziaria.*

L'obiettivo è quindi utilizzare tali sistemi quali supporto all'attività giudiziaria, non demandando un'attività così delicata – e con risvolti su diversi diritti fondamentali dell'individuo – ad un sistema automatico.

La tutela nei confronti dei cittadini emerge in particolar modo dal passaggio *“è sempre riservata al Magistrato la decisione sulla interpretazione della legge, sulla valutazione dei fatti e delle prove e sulla adozione di ogni provvedimento”* che chiarisce come sia sempre il soggetto umano, i cui ragionamenti sono più facilmente comprensibili e trasparenti, a dare una propria interpretazione della norma, con l'Intelligenza artificiale che giocherà un ruolo meramente strumentale di “aiuto” al magistrato.

In ogni caso, il disegno di legge chiarisce altresì che sarà Ministero della giustizia a regolare l'impiego dei sistemi di Intelligenza artificiale da parte degli uffici giudiziari.

5. Il rapporto tra Intelligenza artificiale con la disciplina penalistica e la tutela della proprietà intellettuale

Un elemento fondamentale che non può essere ignorato nel disciplinare l'uso di sistemi tecnologici avanzati sono i possibili risvolti penalistici connessi all'uso di tali strumenti.

Sotto questo punto di vista, in relazione all'Intelligenza artificiale si registra un atteggiamento innovativo del legislatore.

Difatti, se con precedenti innovazioni tecnologiche la scelta del legislatore è stata adattare – per quanto possibile – le figure di reato già codificate (12) alle nuove fattispecie astrattamente rese possibili dai nuovi sistemi tecnologici.

Tuttavia, il ddl non adotta il medesimo approccio, bensì mira a salvaguardare l'uso consapevole dei sistemi di Intelligenza artificiale non già introducendo nuove figure di reato (tranne in un unico caso di cui si dirà a breve), bensì prevedendo che l'uso di sistemi di Intelligenza artificiale costituisca una mera circostanza aggravante.

Ciò comporta che, da un lato il cittadino, che pur potrebbe non comprendere appieno che un determinato uso dell'Intelligenza artificiale possa costituire illecito penalmente rilevante, avrà i medesimi divieti previsti dall'ordinamento precedente e; dall'altro lato, sarà esposto in modo maggiore in termini di severità della pena per un uso contrario alle norme degli strumenti di Intelligenza artificiale.

(11) In particolare, la giurisprudenza civile ed amministrativa si è pronunciata con riguardo alle assegnazioni algoritmiche delle sedi di servizio agli insegnanti, dove è emerso che l'uso dell'algoritmo ha portato ad esiti illogici, irragionevoli e discriminatori. Cfr. Cons. Stato, Sez. VI, 8 aprile 2029, n. 2270.

(12) Si pensi ad esempio al reato di accesso abusivo a sistema informatico, che ha moltissimi elementi in comune con il reato di violazione di domicilio.

Ciononostante, il ddl propone l'introduzione di una nuova fattispecie di reato rubricata nell'art. 612-*quater* che punisce con la reclusione da 1 a 5 anni chiunque “*cagion[i] ad altri un danno ingiusto mediante invio, consegna, cessione, pubblicazione o, comunque, diffusione di immagini o video di persone o cose, ovvero voci o suoni, in tutto o in parte, falsi, generati o manipolati mediante l'impiego di sistemi di intelligenza artificiale, atti ad indurre in inganno sulla loro genuinità o provenienza, è punito. Il delitto è punibile a querela della persona offesa*”.

Il danno ingiusto, pertanto, si verifica (e cioè la condotta assume rilevanza penale) quando vi è l'invio o in generale la diffusione, con qualsiasi forma, l'invio, di persone e di cose, ovvero di voci o suoni, che sono però – e qui emerge la peculiarità di quanto disposto dal legislatore – in tutto o in parte, falsi e quindi *generati o manipolati mediante l'impiego di sistemi di intelligenza artificiale, atti ad indurre in inganno sulla loro genuinità o provenienza*.

Ne consegue che il reato e la relativa punibilità saranno configurabili in caso di diffusione di un prodotto generato con Intelligenza artificiale – e che il prodotto così generato sia suscettibile di indurre un soggetto terzo in errore in merito alla sua genuinità o provenienza.

Oltre ad aggiornare la disciplina penalistica, il legislatore si occupa, però, anche di tutela del diritto di autore, riconfermando nuovamente l'approccio antropocentrico già visto con riferimento ad altri settori.

Il ddl all'articolo 24, infatti integra l'art. 1 della legge n. 633/1941, ricomprendendo nelle “*opere dell'ingegno umano*” anche quelle opere “*realizzate con l'ausilio strumenti di intelligenza artificiale purché il contributo umano sia creativo, rilevante e dimostrabile*”.

Il nuovo testo, pertanto, mira a valorizzare il contributo umano relativamente all'output generato dall'Intelligenza artificiale. Ne consegue che, nell'ottica di valutare la possibilità di proteggere un'opera dell'ingegno creata attraverso strumenti di Intelligenza artificiale, sarà fondamentale la quantificazione e/o qualificazione in concreto, caso per caso, del contributo umano nel processo di realizzazione dell'opera digitale.

Allo stato, però, si deve segnalare come il ddl sia alquanto generico e non individui le modalità per individuare il “*quantum*” dell'apporto umano che rende tutelabile la creazione digitale dal diritto d'autore, ciò che aumenta il rischio di minori tutele per le opere digitali generate con ausilio dell'IA.

Si rileva sul punto, infine, che vi è un forte dibattito in merito all'opportunità di inserire un ulteriore paragrafo – 10-*bis* – all'art. 2 della legge n. 633/1941.

Tale integrazione mirerebbe a dare un'interpretazione autentica di quelli che sono i requisiti di protezione delle opere dell'ingegno, che allo stato sono: i) non particolarmente chiari; ii) oggetto di dibattito internazionale con risultati diversi nei vari paesi del Mondo.

Il *generative IA* ha infatti reso necessario affrontare questo tema in quanto si rischia di ridurre – o aumentare – la protezione per opere dell'ingegno create senza intervento umano (13).

L'Italia, seppur non presenti precedenti giurisprudenziali rilevanti, ha la possibilità con il ddl di fare da capostipite tra gli Stati membri per regolamentare il fenomeno ed evitare che opere dell'ingegno il cui *quid pluris* umano sia pressoché nullo ottengano il riconoscimento delle tutele proprie della legge n. 633/1941.

6. Il concetto di pubblica amministrazione come “Casa di Vetro”

17 giugno 1908. Camera dei deputati. Filippo Turati conia la metafora della pubblica amministrazione come “*Casa di Vetro*” (14).

Sono passati più di cento anni, nonostante ciò, le nuove frontiere digitali e lo sviluppo dell'Intelligenza artificiale riportano ad un'immagine che dovrebbe ancora imporsi quale strumento a presidio di garanzia e democrazia.

Turati ha così anticipato gli esiti a cui è giunta la Costituzione italiana nel 1948. Da allora, l'idea di un agire amministrativo trasparente, intellegibile e corretto, funzionale a colmare la sproporzione tra *pars pubblica* e *pars privata*, tra governanti e governati, ha sempre orientato i giudici amministrativi – ma non solo –, quando, ad esempio, si sono trovati a dover risolvere la questione relativa all'applicabilità dell'accesso civico generalizzato al settore dei contratti pubblici (15).

(13) Il tema è oggetto di grande dibattito a livello globale. Si pensi che negli Usa, la giurisprudenza si è opposta al riconoscimento della tutela per le opere create tramite Intelligenza artificiale; mentre in Cina la posizione è diametralmente opposta, con il riconoscimento di opere create con Intelligenza artificiale purché sia verificabile in concreto l'apporto umano.

(14) F. Turati, *Atti del Parlamento italiano*, Camera dei deputati, sessione 1904-1908, 17 giugno 1908: “*Viene poi l'offesa al decoro della Amministrazione. Che cosa può essere questo decoro? [...] Per mantenere il decoro della Amministrazione, non bisogna non dir niente [...] Io dico che bisognerebbe intanto definire ossia limitare ciò che è segreto di ufficio. Dove un superiore, pubblico interesse non imponga un segreto momentaneo, la casa dell'Amministrazione dovrebb'essere di vetro*”.

(15) La terza sezione del Consiglio di Stato aveva posto, con ord. n. 8501/2019, tra le altre, anche la seguente questione: “*se la disciplina dell'accesso civico generalizzato, di cui al d.lgs. n. 33/2013, come modificato dal d.lgs. n. 97/2016, sia applicabile, in tutto o in parte, in relazione ai documenti relativi alle attività delle amministrazioni disciplinate dal codice dei contratti pubblici di lavori, servizi e forniture, inerenti al procedimento di evidenza pubblica e alla successiva fase esecutiva, ferme restando le limitazioni ed esclusioni oggettive previste dallo stesso codice*.” Con sentenza del 2 aprile 2020, n.

Altresì, va ricordato il fatto che, al fine di garantire un agire pubblico accessibile e trasparente, si è sempre resa necessaria una generalizzata semplificazione dei procedimenti, non potendo esistere trasparenza senza semplificazione.

Ora, trasparenza e semplificazione, da sempre considerate alleate, sono, forse per la prima volta, messe una al cospetto dell'altra. Invero, l'impiego di procedure automatizzate e soluzioni tecnologiche nel ciclo di vita di un provvedimento amministrativo può condurre a sacrificare l'intellegibilità stessa dell'*iter* istruttorio e approvativo.

Fino ad ora, nell'ambito delle procedure di gara sono stati utilizzati per lo più algoritmi non di apprendimento, per il confronto automatico di alcuni parametri caratterizzanti le offerte e pertanto perfettamente conoscibili; tuttavia, non si può escludere che, a breve, saranno disponibili algoritmi di apprendimento molto più complessi e, almeno in parte, imprevedibili negli esiti.

Ecco che, allora, al concetto di "Casa di Vetro" si oppone il concetto di "scatola nera" (o "*black box*" in inglese).

Già dal semplice confronto di queste due espressioni terminologiche, chiaramente antitetiche, è possibile cogliere il paradosso che concerne l'insieme: la metafora della "Casa di Vetro", come anticipato in precedenza, è stata coniata nel 1908 da Filippo Turati; diversamente, l'espressione "scatola nera" non è riconducibile ad uno specifico autore, ma è emersa nel XX secolo, nel contesto delle scienze dei sistemi e dell'ingegneria, per descrivere un sistema o un dispositivo di cui si conoscono gli *input* e gli *output* ma di cui i processi interni rimangono sconosciuti o opachi.

7. Il quadro normativo italiano in materia automazione e digitalizzazione

La prima volta in cui il legislatore italiano si è occupato di disciplinare l'adozione da parte della pubblica amministrazione di atti elaborati mediante strumenti informatici risale al 1997, quando all'art. 15 della l. 15 maggio 1997, n. 59, ha previsto che *"gli atti, dati e documenti formati dalla pubblica amministrazione e dai privati con strumenti informatici o telematici, i contratti stipulati nelle medesime forme, nonché la loro archiviazione e trasmissione con strumenti informatici, sono validi e rilevanti a tutti gli effetti di legge"*.

Tuttavia, la norma è rimasta sostanzialmente priva di applicazione concreta, non essendo mai stati emanati i relativi regolamenti attuativi (16).

Nel 2005 il legislatore è intervenuto nuovamente con un testo organico volto ad orientare la trasformazione digitale della pubblica amministrazione. Si tratta del Codice dell'amministrazione digitale (c.d. Cad), di cui al d.lgs. 7 marzo 2005, n. 82, che all'art. 12 stabilisce che *"le pubbliche amministrazioni nell'organizzare autonomamente la propria attività utilizzano le tecnologie dell'informazione e della comunicazione per la realizzazione degli obiettivi di efficienza, efficacia, economicità, imparzialità, trasparenza, semplificazione e partecipazione nel rispetto dei principi di uguaglianza e di non discriminazione"*.

Successivamente, col Decreto semplificazioni del 2020 è stato modificato l'art. 3-bis della l. 7 agosto 1990, n. 241 (17), nel senso che *"per conseguire maggiore efficienza nella loro attività, le amministrazioni pubbliche agiscono mediante strumenti informatici e telematici, nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati"* (18).

Detto adeguamento normativo deriva, oltre che dall'affermazione dei principi in ambito europeo, anche da una forte spinta della giurisprudenza amministrativa italiana, la quale è giunta a riconoscere piena valenza giuridica e amministrativa alla regola algoritmica. Essa, sia pure declinata in forma matematica, deve dunque *"soggiacere ai principi generali dell'attività amministrativa, quali quelli di pubblicità e trasparenza (art. 11, l. n. 241/1990), di ragionevolezza, di proporzionalità"* (19).

Altresì, mentre le disposizioni sopra richiamate concernono in via generale l'azione amministrativa, nel 2023 il legislatore si è preoccupato di intervenire nel settore specifico della contrattualistica pubblica delegando il

10, il medesimo collegio ha ritenuto *"applicabile l'accesso civico generalizzato ai documenti inerenti ai contratti pubblici, inclusi quelli relativi all'esecuzione dei medesimi"*.

(16) L'art. 15, secondo periodo, della l. 15 maggio 1997, n. 59, dispone che *"i criteri e le modalità di applicazione del presente comma sono stabiliti, per la pubblica amministrazione e per i privati, con specifici regolamenti da emanare entro centottanta giorni dalla data di entrata in vigore della presente legge ai sensi dell'articolo 17, comma 2, della legge 23 agosto 1988, n. 400"*.

(17) Articolo modificato dall'art. 12, c. 1, lett. b), d.l. 16 luglio 2020, n. 76, convertito, con modificazioni, dalla l. 11 settembre 2020, n. 120.

(18) Sul punto, F. Cardarelli, "Artt. 3-3-bis", in M.A. Sandulli (a cura di), *Codice dell'azione amministrativa*, Milano, Giuffrè, 2ª ed., 2017; S. Rossa, *Contributo allo studio delle funzioni amministrative digitali. Il processo di digitalizzazione della Pubblica Amministrazione e il ruolo dei dati aperti*, Cedam-Walters Kluwer, 2021.

(19) In tale senso, Cons. Stato, Sez. VI, n. 2270/2019, cit., successivamente richiamato da Cons. Stato, Sez. VI, 13 dicembre 2020, n. 881.

Governo (20) ad adottare una specifica disposizione che ammettesse la possibilità per le stazioni appaltanti e gli enti concedenti di ricorrere anche ad automatismi nella valutazione delle offerte.

Le indicazioni del legislatore nazionale sono state recepite all'articolo 30 del Nuovo Codice dei contratti pubblici, ex d.lgs. n. 36/2023 (anche semplicemente, Nuovo Codice).

8. *L'Intelligenza artificiale e le nuove soluzioni tecnologiche applicate al ciclo di vita dei contratti pubblici*

L'articolo 30 del Nuovo Codice rappresenta un'assoluta novità per l'ordinamento italiano (21), in quanto per la prima volta si interviene puntualmente per disciplinare l'automazione dell'attività amministrativa, sia pur con riferimento al settore dei contratti pubblici; diversamente, le norme poc'anzi citate si sono occupate solamente di incentivare la diffusione delle nuove tecnologie senza fornire delle direttive operative concrete (22).

In sostanza, la norma conferisce effettività ad una digitalizzazione dell'agire amministrativo che è stata, fino a poco tempo fa, solamente formale (23) (o quasi); di recente, invece, si è iniziato a parlare di digitalizzazione sostanziale poiché *“capace di trasformare profondamente i processi decisionali, mutarne le direttrici di fondo e sostituire la macchina e la tecnologia alla componente umana”* (24).

Il primo comma prescrive alle stazioni appalti e agli enti concedenti, ove possibile, al fine di migliorare l'efficienza, di automatizzare le proprie attività ricorrendo a soluzioni tecnologiche, ivi incluse l'Intelligenza artificiale e le tecnologie di registri distribuiti.

Ciò deve avvenire, oltre che nel rispetto delle specifiche disposizioni in materia, anche nel rispetto dei principi di: *“a) conoscibilità e comprensibilità, per cui ogni operatore economico ha diritto a conoscere l'esistenza di processi decisionali automatizzati che lo riguardano e, in tal caso, a ricevere informazioni significative sulla logica utilizzata; b) non esclusività della decisione algoritmica, per cui comunque esiste nel processo decisionale un contributo umano capace di controllare, validare ovvero smentire la decisione automatizzata; c) non discriminazione algoritmica, per cui il titolare mette in atto misure tecniche e organizzative adeguate al fine di impedire effetti discriminatori nei confronti degli operatori economici”* (25).

Con riferimento alla necessità che siano rispettati i principi di “conoscibilità e comprensibilità”, risulta opportuno partire dalle riflessioni della giurisprudenza amministrativa antecedente al Nuovo Codice, che si è espressa nel senso di ritenere la regola tecnica che governa ciascun algoritmo come una regola amministrativa generale, costruita dall'uomo e non dalla macchina, per poi essere (solo) applicata a quest'ultima.

La regola algoritmica, infatti, *“non può lasciare spazi applicativi discrezionali (di cui l'elaboratore elettronico è privo), ma deve prevedere con ragionevolezza una soluzione definita per tutti i casi possibili, anche i più improbabili (e ciò la rende in parte diversa da molte regole amministrative generali)”* (26).

In altri termini, l'Amministrazione deve sempre compiere un ruolo *ex ante* di mediazione e composizione di interessi, a maggior ragione nel caso di apprendimento progressivo e di *deep learning*, non potendo la discrezionalità amministrativa essere demandata al *software*.

Le risultanze della giurisprudenza amministrativa sono confluite nella disposizione in commento, atteso che al secondo comma, è previsto che le stazioni appaltanti e gli enti concedenti, in sede di acquisto o di sviluppo delle soluzioni tecnologiche, ivi incluse l'Intelligenza artificiale e le tecnologie di registri distribuiti, assicurino: (i) la disponibilità del codice sorgente, della relativa documentazione, nonché di ogni altro elemento utile a comprenderne le logiche di funzionamento; (ii) l'introduzione negli atti di indizione delle gare di clausole volte ad

(20) Delega al Governo in materia di contratti pubblici, G.U., Serie generale, 24 giugno 2022, n. 146, lett. t): *“individuazione delle ipotesi in cui le stazioni appaltanti possono ricorrere ad automatismi nella valutazione delle offerte e tipizzazione dei casi in cui le stazioni appaltanti possono ricorrere, ai fini dell'aggiudicazione, al solo criterio del prezzo o del costo, con possibilità di escludere, per i contratti che non abbiano carattere transfrontaliero, le offerte anomale determinate sulla base di meccanismi e metodi matematici, tenendo conto anche della specificità dei contratti nel settore dei beni culturali e prevedendo in ogni caso che i costi della manodopera e della sicurezza siano sempre scorporati dagli importi assoggettati a ribasso”*.

(21) In un'ottica comparatistica, v. A. Masucci, *L'algoritmizzazione delle decisioni amministrative tra Regolamento europeo e leggi degli Stati membri*, in *Dir. pubbl.*, 2020, 943 ss.

(22) R. Calvara, *L'impiego di sistemi di Intelligenza artificiale nei contratti pubblici: l'alba di un nuovo potere amministrativo?*, in *Intelligenza artificiale e pubbliche amministrazioni: una tripla visione in chiave comparativa*, Atti del X Congresso della Rete internazionale del diritto europeo (Ride)-Digitapia, Barcellona, Universitat Oberta de Catalunya, 5 dicembre 2023.

(23) Il tema dell'utilizzo della tecnologia nel contesto dell'agire amministrativo è già stato trattato nel 1979 da M.S. Giannini in un rapporto trasmesso alle Camere avente ad oggetto i *“principali problemi della amministrazione dello Stato”*. Successivamente, anche A. Predieri, *Gli elaboratori elettronici nell'amministrazione dello Stato*, Bologna, il Mulino, 1971; S. Rodotà, *Tecnologie e diritti*, 1995, Bologna, il Mulino, 2021.

(24) B.G. Mattarella, M. Ramajoli (diretto da), *Enciclopedia del diritto: funzioni amministrative*, Milano, Giuffrè, 2022, 77.

(25) Art. 30, c. 3, del Nuovo codice.

(26) In tale senso, Cons. Stato, Sez. VI, 8 aprile 2019, n. 2270; Tar Puglia-Bari, Sez. III, 22 aprile 2022, n. 552.

assicurare le prestazioni di assistenza e manutenzione necessarie alla correzione degli errori e degli effetti indesiderati derivanti dall'automazione.

A detto intervento necessario dell'Amministrazione a monte, si combina l'onere per la stessa di pubblicare l'elenco delle soluzioni tecnologiche utilizzate nello svolgimento della propria attività sul sito istituzionale, nella sezione "*Amministrazione trasparente*", a norma del quinto comma dell'articolo 30 del Nuovo Codice (27).

Ciononostante, il dubbio che residua riguarda sempre la sufficienza di detti oneri posti in capo all'Amministrazione a fronte del (prossimo) futuro sviluppo di meccanismi algoritmici di cui non sarà possibile descrivere puntualmente le modalità operative né, tanto meno, prevedere le risultanze finali.

Pur volendo considerare l'onere di motivazione che grava sulla Pubblica amministrazione ai sensi dell'art. 3 della Legge n. 241/1990 – anche al fine di assicurare il pieno esercizio del diritto di difesa ai sensi degli articoli 24 e 113 della Costituzione – non sembra potersi scongiurare *in toto* il rischio di una mancata piena conoscibilità e prevedibilità dei meccanismi di Intelligenza artificiale e del *machine learning*, atteso che il cortocircuito interviene in un momento antecedente a quello finale di adozione del provvedimento amministrativo.

Come potrebbe, infatti, esplicarsi pienamente l'onere motivazionale dell'Amministrazione rispetto ad un algoritmo che evolve in autonomia e che si auto-alimenta?

Altresì, sempre il terzo comma dell'articolo 30 – coerentemente con l'approccio seguito dal legislatore comunitario – ha previsto quella che in dottrina è stata definita "riserva di umanità" (28), in virtù della quale esiste sempre nel processo decisionale l'intervento umano volto a controllare, validare e modificare la decisione automatizzata.

Ciò dovrebbe essere "*presidio per una intelligenza artificiale al servizio dell'amministrazione più giusta, efficace ed equa [...] [nonché] un fondamentale fattore di legittimazione sociale dell'automazione avanzata a mezzo di intelligenza artificiale*" (29).

In sostanza, vi è un intervento umano successivo rispetto al momento di scelta "autonoma" del software, al fine di avere quale esito ultimo una scelta "*assistita, ovvero assunta dall'uomo ma con il supporto dello strumento dell'intelligenza artificiale*" (30).

Tuttavia, è estremamente delicato il tema della perimetrazione dell'attività del soggetto preposto. In primo luogo, perché sono richieste delle competenze tecniche sofisticate, oltre alla capacità di interpretare e dare attuazione alle risultanze del sistema algoritmico utilizzato (31); in secondo luogo, perché sono individuabili vari livelli di controllo corrispondenti ad un ruolo crescente della componente umana.

Con riferimento a questo secondo aspetto, si può prospettare una totale assenza di sorveglianza e dominio nei contesti completamente automatizzati (*human out of the loop*), ovvero un controllo limitato alla fase di sviluppo e funzionamento con possibilità per il "sorvegliante" di interrompere e modificare il lavoro (*human in the loop*) o, ancora, il controllo pieno del funzionamento in un contesto di automazione estremamente limitato (*human in command*) (32).

In ogni caso, questa disposizione assume rilevanza rispetto all'assunzione di responsabilità da parte dell'amministrazione procedente nei confronti dei destinatari del provvedimento amministrativo che si ritengono lesi nelle loro posizioni giuridiche soggettive.

Infatti, se da un lato il giudice deve valutare la correttezza del processo automatizzato in tutte le sue componenti, dovendosi considerare il software a tutti gli effetti come un "*atto amministrativo informatico*" (33); dall'altro lato, l'interessato ha diritto a non essere sottoposto ad una decisione fondata unicamente sul trattamento automatizzato ove la stessa "*produca effetti giuridici che lo riguardano o incida in modo analogo significativamente sulla sua persona*" (34).

In questa ottica, il giudice deve sempre essere reso edotto dei criteri ideati dal programmatore del software per lo sviluppo dell'algoritmo. Questi, pertanto, devono essere logici, ragionevoli, non discriminatori e sempre

(27) Come si legge nella *Relazione illustrativa al Nuovo Codice dei contratti pubblici*.

(28) Si legga in tal senso, A. Corrado, *I nuovi contratti pubblici, intelligenza artificiale e blockchain: le sfide del prossimo futuro*, in <www.federalismi.it>, 26 luglio 2023; G. Gallone, *Riserva di umanità e funzioni amministrative*, Padova, Cedam, 2023.

(29) G. Gallone, *op. cit.*

(30) *Ibidem*.

(31) In questo senso, S. Foà, *Intelligenza artificiale e cultura della trasparenza amministrativa. Dalle "scatole nere" alla "casa di vetro"?*, in *Dir. amm.*, 2023, 515. In particolare, l'autore evidenzia come il soggetto preposto debba essere in grado di "*comprendere appieno le capacità e i limiti del sistema di IA ad alto rischio ed essere in grado di monitorarne debitamente il funzionamento, in modo che i segnali di anomalie, disfunzioni e prestazioni inattese possano essere individuati e affrontati quanto prima*".

(32) L. Rinaldi, *Intelligenza artificiale, diritti e doveri nella Costituzione italiana*, in *DPCE online*, 2022, 201.

(33) In questo senso, Cons. Stato, Sez. VI, 8 aprile 2019, n. 2270; Tar Puglia, Bari, Sez. III, 22 aprile 2022, n. 552.

(34) Il riferimento è all'art. 22 Gdpr, Regolamento europeo sulla protezione dei dati personali.

accessibili al destinatario dell'azione amministrativa, in modo tale da consentire la piena sindacabilità in sede giurisdizionale.

Da ultimo, con riferimento al principio di non discriminazione algoritmica, è doveroso il riferimento al c. 4 dell'art. 30, a norma del quale le stazioni appaltanti e gli enti concedenti adottano ogni misura tecnica e organizzativa atta a garantire che siano rettificati i fattori che comportano inesattezze dei dati e che sia minimizzato il rischio di errori. Ciò, al fine di impedire effetti discriminatori nei confronti di persone fisiche sulla base della nazionalità, dell'origine etnica, delle opinioni politiche, della religione, delle convinzioni personali, dell'appartenenza sindacale, dei caratteri somatici, dello *status* genetico, dello stato di salute, del genere o dell'orientamento sessuale.

Dall'esigenza del legislatore di specificare quali possano essere in concreto gli effetti discriminatori derivanti dall'utilizzo di sistemi intelligenti e automatizzati, è possibile desumere tutta la pericolosità insita negli stessi e nella loro progressiva implementazione.

In particolare, il rischio è rappresentato dalla possibilità che sorgano dei *bias* (35), ossia dei veri e propri pregiudizi, che vengano successivamente reiterati dalla formula algoritmica.

In Italia, ad esempio, la giurisprudenza si è trovata di fronte a discriminazioni determinate dalla regola algoritmica nelle procedure di assunzione del personale docente e di scorrimento delle graduatorie (36). Nel caso di specie, il giudice amministrativo si è limitato a dichiarare l'inammissibilità del ricorso per difetto di giurisdizione. Occorre, dunque, evidenziare come alle volte la stessa giurisprudenza fatichi ad esprimersi rispetto a fattispecie in cui sono coinvolte tecnologie innovative.

Nell'ordine del giorno si ritrova spesso anche il tema dell'opportunità o meno di impiegare sistemi di riconoscimento facciale idonei a collezionare dati biometrici. Anche in questo caso, da un lato l'utilizzo di sistemi di Intelligenza artificiale potrebbe rivelarsi utile quale strumento di contrasto delle minacce all'ordine pubblico, dall'altro lato, si potrebbero immagazzinare dati sensibili come convinzioni politiche o religiose.

Ciononostante, la medesima giurisprudenza si è dimostrata aperta all'impiego delle nuove tecnologie, ivi inclusi algoritmi ed Intelligenza artificiale, sia pur con delle riserve: in primo luogo, deve sempre esistere una decisione umana al termine del processo (torna qui la già menzionata "riserva di umanità"); in secondo luogo, deve sempre valere il "*principio del servizio*", in virtù del quale l'Intelligenza artificiale deve essere strumentale e non sostitutiva dell'opera umana. In altri termini, anche ai fini dell'imputazione delle responsabilità connesse all'esercizio del potere amministrativo, l'Intelligenza artificiale deve essere concepita quale *quid pluris* a servizio dell'uomo.

D'ogni modo, sono varie le tematiche che restano allo stato insolute: dalla possibile futura sostituzione integrale dell'opera umana con tecnologie che si auto-alimentano; alla eventuale sindacabilità anche *ex post* del corretto esercizio del potere amministrativo da parte del cosiddetto "Giudice Robot".

È chiaro che, in questi casi, sarebbero violati il principio della riserva di umanità, quello del servizio e, soprattutto, il principio di trasparenza.

9. Bilanciamento dei principi sottesi all'agire amministrativo. Una nuova interpretazione costituzionalmente orientata

L'attività amministrativa deve rispondere ad uno stretto principio di legalità, assicurando al contempo economicità, efficacia, imparzialità, pubblicità e trasparenza. Peraltro, i rapporti tra il cittadino e la pubblica amministrazione devono essere improntati ai principi della collaborazione e della buona fede (37).

A detti principi di portata generale, il legislatore nel 2023 ha affiancato ulteriori principi volti a governare la materia della contrattualistica pubblica. Primo fra tutti, il principio del risultato – che non casualmente è stato collocato al primo articolo del Nuovo Codice – in virtù del quale "*le stazioni appaltanti e gli enti concedenti perseguono il risultato dell'affidamento del contratto e della sua esecuzione con la massima tempestività e il migliore rapporto possibile tra qualità e prezzo, nel rispetto dei principi di legalità, trasparenza e concorrenza*".

Il principio del risultato rappresenta una derivazione evoluta del più generale principio di buon andamento di cui all'art. 97 Cost. e si colloca nel contesto della legalità e della concorrenza. La sua codificazione è finalizzata

(35) Per comprendere meglio questo fenomeno, è possibile menzionare il caso Optum, un sistema utilizzato negli ospedali statunitensi per prevedere quali pazienti avrebbero potuto avere necessità di cure mediche ulteriori. Detto algoritmo, tuttavia, avrebbe progressivamente iniziato a discriminare i pazienti neri rispetto a quelli bianchi, seppur il colore della pelle non rientrasse tra le variabili considerate *ab origine*. Invero, il *bias* avrebbe trovato fondamento nel dato per cui i pazienti di colore risultavano aver sostenuto in media spese sanitarie inferiori rispetto ai pazienti bianchi.

(36) Un caso ha riguardato la richiesta di annullamento di una serie di provvedimenti inerenti a una procedura di assegnazione delle sedi scolastiche presso cui effettuare attività di supplenza. Tuttavia, il Tar Lazio ha dichiarato inammissibile il ricorso per difetto di giurisdizione sottolineando semplicemente che "*l'algoritmo, ossia il software, deve essere considerato a tutti gli effetti come un 'atto amministrativo informatico', posto che tale affermazione risulta essere condivisibile solo laddove l'utilizzo del mezzo informatico sia strumentale all'esercizio di un potere autoritativo di stampo pubblicistico, su cui sussiste la giurisdizione del giudice amministrativo*" (cfr. Tar Lazio, Roma, Sez. III-bis, 16 maggio 2023, n. 8384).

(37) Come dispone l'art. 1 della l. n. 241/1990.

a dare attuazione, nel settore dei contratti pubblici, ai principi di efficienza, efficacia ed economicità ed “è *perseguito nell'interesse della comunità e per il raggiungimento degli obiettivi dell'Ue*” (38).

In sostanza, deve essere perseguito non “*un risultato purché sia*” ma un risultato “virtuoso” (39).

Interessante è poi il riferimento alla trasparenza al secondo comma dell'art. 1 del Nuovo Codice, considerata funzionale alla massima semplicità e celerità nella corretta applicazione del medesimo decreto legislativo.

Analogo riferimento alla trasparenza attua anche l'art. 2 del Nuovo Codice che ha positivizzato il principio della reciproca fiducia nell'azione legittima, trasparente e corretta dell'amministrazione, dei suoi funzionari e degli operatori economici.

La fiducia rappresenta quindi il presupposto per l'attribuzione di potere alla pubblica amministrazione, specie se di natura discrezionale: invero, ogni conferimento di potere presuppone “*la fiducia dell'ordinamento giuridico verso l'organo destinatario dell'attribuzione: esplicitare a livello normativo questo presupposto culturale e giuridico promuove il senso di appartenenza dell'Amministrazione allo Stato-comunità, scongiora l'inerzia, valorizza la capacità e orienta verso il rispetto della legalità sostanziale*” (40).

In definitiva, affinché l'azione della pubblica amministrazione possa considerarsi rispettosa del principio di legalità sostanziale, deve essere conoscibile e, successivamente, eventualmente sindacabile nella sua correttezza.

Il principio di trasparenza, che si esprime proprio nel concetto di conoscibilità, rappresenta il fondamento della democrazia amministrativa in uno Stato di diritto e “*costituisce anche un caposaldo del principio di buon funzionamento della pubblica amministrazione, quale “casa di vetro” improntata ad imparzialità, intesa non quale mera conoscibilità, garantita dalla pubblicità, ma anche come intelligibilità dei processi decisionali e assenza di corruzione*” (41).

Altresì, il diritto dei cittadini ad accedere ai dati in possesso della pubblica amministrazione risponde, oltre che a principi di pubblicità e trasparenza, anche al buon funzionamento della pubblica amministrazione ai sensi dell'art. 97 Cost., consentendo la partecipazione diffusa di tutti i cittadini alla cura della “cosa pubblica” (42) ai sensi degli artt. 1 e 2 Cost.

È infatti necessario mettere il singolo nella condizione di poter adempiere al dettato dell'art. 2 Cost., in virtù del quale la Repubblica, oltre a riconoscere e garantire i diritti inviolabili dell'uomo, richiede ad ognuno l'adempimento dei doveri inderogabili di solidarietà politica, economica e sociale. Come ricorda il giudice costituzionale: “*non c'è organizzazione che, direttamente o almeno indirettamente, non sia finalizzata a diritti, così come non c'è diritto a prestazione che non condizioni l'organizzazione*” (43).

10. Possibili evoluzioni future delle nuove tecnologie nel settore dei contratti pubblici

Dalla realtà fattuale odierna e con specifico riferimento al settore della contrattualistica pubblica, è possibile cogliere qualche spunto di riflessione e provare ad immaginare quali possano essere le evoluzioni tecnologiche del prossimo futuro.

Ad esempio, si può pensare alla creazione di un “assistente virtuale del Rup”, ossia un percorso di supporto alle attività che segnano il ciclo di vita del contratto, dalla fase di programmazione a quella di esecuzione.

Oppure, si può immaginare la realizzazione di meccanismi di controllo automatizzati che siano posti a supporto della Commissione di gara.

Un esempio potrebbe essere un sistema automatizzato che verifichi la detenzione e la regolarità delle certificazioni presentate dagli operatori economici concorrenti; ovvero l'impiego di procedure automatizzate nel contesto della valutazione dell'offerta tecnica ed economica.

Con riferimento a questo secondo aspetto, potrebbero riscontrarsi alcuni vantaggi sia in termini di tempistiche che in termini di oggettività della valutazione, non potendosi al contempo escludere l'esistenza di qualche criticità. Da un lato, infatti, si potrebbe scongiurare il rischio di un'interpretazione differenziata caso per caso di clausole ambigue, con una conseguente maggiore controllabilità in sede giudiziaria; dall'altro lato, è vero anche che, spesso e volentieri, è necessaria una differenziazione nell'interpretazione, atteso che ogni fattispecie concreta ha dei profili che la rendono dissimile da ogni altra.

(38) Art. 1, c. 3, del Nuovo Codice.

(39) Si legge nella *Relazione illustrativa* al Nuovo Codice.

(40) In questo senso si esprime la *Relazione illustrativa* al Nuovo Codice.

(41) Cons. Stato, Ad. plen., 2 aprile 2020, n. 10.

(42) In questo senso, Cons. Stato, Ad. plen., 12 aprile 2016, n. 7; Sez. III, 2 settembre 2014, n. 4460: “*La luce della trasparenza feconda il seme della conoscenza tra i cittadini e concorre, da un lato, al buon funzionamento della pubblica amministrazione ma, dall'altro, anche al soddisfacimento dei diritti fondamentali della persona, se è vero che organizzazione amministrativa e diritti fondamentali sono strettamente interrelati*”.

(43) Cfr. Corte cost., 27 novembre 1998, n. 383.

Ancora, non è difficile immaginare la creazione di sistemi di Intelligenza artificiale volti a realizzare indagini preliminari di mercato finalizzati ad individuare con precisione l'oggetto del contratto nonché la misura in cui esso è presente sul mercato.

Da ultimo, si può considerare la figura del cosiddetto “*smart contract*”, e cioè un insieme di protocolli informatici volti ad imporre il rispetto, tramite controlli e verifiche, della negoziazione e dell'esecuzione di un contratto.

11. Considerazioni conclusive

In questa sede, si potrebbero mettere sul tavolo una molteplicità di spunti di riflessione e avanzare altrettante predizioni rispetto a quella che sarà, più o meno a breve, l'evoluzione dell'Intelligenza artificiale in Italia, prima ancora che in Europa e nel mondo.

Senza dubbio, siamo in grado, almeno parzialmente, di immaginare quelle che possono essere le risultanze più probabili rispetto alla direzione in cui si stanno muovendo sia il legislatore che la giurisprudenza.

La verità, però, è che prefigurare oggi ciò che sarà, verosimilmente, domani, potrebbe apparire per certi versi come un atto di presunzione. Presunzione, perché a fronte di strumenti “intelligenti” e automatizzati che possono trovare legittimazione nei loro algoritmi e nelle loro formule matematiche, noi siamo solo umani.

Non siamo intelligenze generative. Non siamo nemmeno intelligenze predittive. Non siamo infallibili calcolatori.

Da ciò, derivano tanto il rischio di peccare di umiltà, quanto la difficoltà di “ingabbiare” all'interno di categorie giuridiche un fenomeno così magmatico e poliforme come è quello dell'Intelligenza artificiale.

D'altro canto, però, non possiamo nemmeno rinunciare in partenza a regolare la complessità del tutto (44).

Forse, esiste una terza via. Forse, si può partire dal cercare l'*aurea mediocritas* oraziana, ossia un equilibrio tra le due forze: macchina e uomo.

Forse, consapevoli della nostra brevità e limitatezza – come scrisse Giacomo Leopardi “*umana cosa picciol tempo dura*” (45) – potremmo certamente impiegare l'Intelligenza artificiale come strumento di velocizzazione ed efficienza, sempre mantenendo la componente umana quale presidio di libertà, diritti e garanzie.

Per fare ciò e fugare il rischio della “*black box*”, si deve tenere a mente la metafora turatiana in modo da costruire, passo dopo passo, la nostra “Casa di Vetro”.

* * *

(44) Qui, si può riprendere l'eterno confronto tra l'idea di Isaac Asimov – autore delle Tre Leggi della robotica negli anni '40 – e i suoi contestatori: mentre, secondo Asimov una macchina progettata bene non poteva presentare alcun rischio; diversi autori successivi ritennero che non fosse possibile limitare con certezza il comportamento dei robot inventando ed applicando un insieme di regole.

(45) In G. Leopardi, *Canti* (Leopardi-Donati), XLI. Dello stesso.

INTELLIGENZA ARTIFICIALE E RESPONSABILITÀ AMMINISTRATIVO-CONTABILE

di Paolo Evangelista (*)

Abstract: Nell'ordinamento giuridico italiano diverse sono le fonti normative che disciplinano l'utilizzo dell'Intelligenza Artificiale in tutto l'agire della pubblica amministrazione. La giurisprudenza più recente del giudice amministrativo ha peraltro riconosciuto la possibilità del ricorso all'IA non solo nelle ipotesi di attività amministrativa vincolata ma anche, entro certi limiti, nell'esercizio di attività discrezionale. L'obiettivo del lavoro, una volta configurata una spesa illegittima sostenuta dalla Pubblica amministrazione a seguito di non corretto utilizzo dell'IA, è quello di verificare se sia o meno possibile imputare una responsabilità amministrativo-contabile per danno erariale in presenza dei requisiti del nesso di causalità e del c.d. elemento psicologico del dolo o della colpa grave.

In the Italian legal system, there are several sources of legislation governing the use of Artificial Intelligence (AI) throughout the actions of the public administration. The most recent case-law of the Administrative Court has, however, recognised the possibility of recourse to AI not only in cases of restricted administrative activity but also, within certain limits, in exercising discretionary activities. The main goal of the work, once an illegitimate claim has been made by the public administration following incorrect use of AI, is to verify whether or not it is possible to charge an administrative-accounting liability for fiscal damage considering the requirements of causality and the so-called psychological element of deliberate intention and gross negligence.

Sommario: 1. Considerazioni preliminari. – 2. Le fonti normative che disciplinano l'utilizzo dell'IA nell'agire della P.A. – 3. L'utilizzo dell'IA nella giurisprudenza del giudice amministrativo T.A.R./Consiglio di Stato. – 4. IA ed i presupposti per configurare una responsabilità amministrativa-contabile. – 4.1. Ipotesi di danno erariale. C.d. indiretto. – 4.2. IA e danno erariale diretto da attività amministrativa/provvedimentale illegittima. – 5. Conclusioni.

1. Considerazioni preliminari

Il quesito di fondo a cui si è chiamati a fornire una risposta con il presente lavoro è se sia possibile o meno imputare ad un dipendente pubblico una responsabilità amministrativo-contabile per danno erariale conseguente ad applicazione/utilizzo dell'Intelligenza Artificiale nello svolgimento dell'attività amministrativa o nell'erogazione di un servizio pubblico.

L'individuazione della risposta al predetto quesito non è agevole perché in caso di automazione 'avanzata' a mezzo di intelligenza artificiale la dominabilità dei suoi risultati, da parte dell'uomo, come sappiamo è assai ridotta.

Al riguardo occorre innanzitutto precisare cosa si intende per sistema di IA.

Nel recente Regolamento 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 (art. 3 – Definizioni)

Il 'sistema di IA' viene così definito: "un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali".

In dottrina è stato anche sottolineato che "per intelligenza artificiale, dunque, possono sinteticamente intendersi tutti quei sistemi e/o procedimenti e/o software che "simulano" l'intelligenza umana, al fine sostanzialmente di porre in essere un dato compito (con risparmio di tempo e maggiore precisione rispetto a quanto potrebbe fare un essere umano)" (1).

Nell'ambito dell'IA, è stato anche precisato quanto segue: "si distinguono il Machine Learning che si occupa dello sviluppo di algoritmi e tecniche finalizzate all'apprendimento automatico mediante la statistica computazionale e l'ottimizzazione matematica e il Deep Learning quale classe di algoritmi di apprendimento automatico che utilizza livelli multipli per estrarre progressivamente caratteristiche di livello superiore dall'input grezzo. Ed invero, il Machine Learning può essere definito come il metodo che "allena" l'intelligenza artificiale, attraverso il quale si permette alle macchine di apprendere in modo che possano poi svolgere un compito o un'attività senza che siano preventivamente programmati, bensì automaticamente; il Deep Learning, invece, permette di emulare la mente dell'uomo in quanto rappresenta un modello di apprendimento ispirato alla struttura ed al funzionamento del cervello. In questo caso, il modello matematico da solo non basta, il Deep Learning necessita di reti neurali artificiali progettate ad hoc (deep artificial neural network) e di una capacità computazionale potente capace di "reggere" differenti strati di calcolo e analisi (che è quello che succede con le connessioni neurali del cervello umano) (2).

(*) P. Evangelista è Presidente di sezione della Corte dei conti.

(1) N.M. Tritto, *Il danno erariale da intelligenza artificiale*, in www.federalismi.it, 20 marzo 2024.

(2) A. Bove, *Digitalizzazione e Intelligenza artificiale nell'attività della Pubblica Amministrazione. Ammissibilità, limiti e condizioni*, in *Gazzetta forense*, novembre-dicembre 2023, p. 871

Nel nostro ordinamento giuridico, come vedremo, il Legislatore ha da tempo previsto il ricorso all'IA per rendere più efficiente ed efficace l'attività in tutti i settori della pubblica amministrazione (3).

Una sperimentazione sull'utilizzo dell'IA è in corso anche in sede di controlli finanziari degli enti territoriali esercitati dalla Sezione regionale di controllo della Corte dei conti (4), al fine di prevedere/prevenire situazioni di dissesto finanziario dell'ente controllato.

2. Le fonti normative che disciplinano l'utilizzo dell'IA nell'agire della P.A.

Cominciamo ad evidenziare che, a livello normativo, con le modifiche apportate dall'art. 12 del d.l. n. 76/2020 (5) l'art. 3 bis della l. n. 241/1990 dispone che *“Per conseguire maggiore efficienza nella loro attività, le amministrazioni pubbliche agiscono mediante strumenti informatici e telematici, nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati.”*; quindi rispetto al previgente testo introdotto con la l. n. 15/2005, non vi è più una disposizione volta semplicemente ad incentivare l'uso della telematica bensì vi è una previsione in termini di obbligo/dovere da parte della pubblica amministrazione alla digitalizzazione ed informatizzazione dei procedimenti amministrativi, nell'ottica di una effettiva semplificazione ed efficienza (6).

È da evidenziare inoltre che tutte le disposizioni della parte II del libro I del nuovo codice dei contratti – dall'art. 19 all'art. 36 – mirano all'obiettivo di fondo di digitalizzare l'intera procedura dei contratti pubblici, fondandola sulla creazione di documenti nativi digitali e sull'utilizzo di piattaforme digitali in modo da rendere possibile l'interazione con le banche dati esistenti (7).

L'art. 30 del nuovo codice dei contratti prevede in particolare che le stazioni appaltanti devono provvedere, ove possibile, *“ad automatizzare le proprie attività ricorrendo a soluzioni tecnologiche, ivi incluse l'intelligenza artificiale”*.

3. L'utilizzo dell'IA nella giurisprudenza del giudice amministrativo T.A.R./Consiglio di Stato

Nella giurisprudenza amministrativa Tar/Consiglio di Stato possiamo inoltre ritenere consolidato il principio secondo cui è possibile il ricorso a procedure automatizzate o all'Intelligenza Artificiale non solo nell'attività vincolata ma, entro certi limiti, anche nell'attività discrezionale della pubblica amministrazione (8).

In particolare, nella sentenza del Consiglio di Stato, Sez. VI, 13 dicembre 2019, n. 8472, è stato affermato che, *“se il ricorso agli strumenti informatici può apparire di più semplice utilizzo in relazione alla cosiddetta attività vincolata, nulla vieta che i [fini stabiliti dalla legge], perseguiti con il ricorso all'algoritmo informatico, possano perseguirsi anche in relazione ad attività connotata da ambiti di discrezionalità. Piuttosto, se nel caso dell'attività vincolata ben più rilevante, sia in termini quantitativi che qualitativi, potrà essere il ricorso a strumenti di automazione della raccolta e valutazione dei dati, anche l'esercizio di attività discrezionale, in specie tecnica, può in astratto beneficiare delle efficienze e, più in generale, dei vantaggi offerti dagli strumenti stessi”*.

È stato al riguardo osservato che soltanto nell'ambito dell'attività discrezionale della pubblica amministrazione può immaginarsi l'impiego di tecnologie informatiche che non si limitano ad eseguire istruzioni preventivamente programmate, ma pervengono all'autonoma formulazione di decisioni in relazione alle caratteristiche della fattispecie concreta. *“Viceversa, nei casi in cui la legge determina analiticamente il contenuto dell'atto da adottare ovvero prescrive lo svolgimento di valutazioni di ordine tecnico-scientifico, le capacità di autonomia decisionale dei moderni programmi di intelligenza artificiale non possono essere utilmente sfruttate”* (9).

In una decisione del Tar Campania (Sez. III, 14 novembre 2022, n. 7003) dove vengono richiamate pronunce della sez. VI del Consiglio di Stato tra cui la n. 2270/2019, viene precisato che *“Il ricorso all'algoritmo [...] anche in funzione parzialmente decisionale nei procedimenti a basso tasso di discrezionalità, non può mai comportare un abbassamento del livello delle tutele garantite dalla legge sul procedimento amministrativo, ed in particolare di quelle sulla individuazione del responsabile del procedimento, sull'obbligo di motivazione, sulle garanzie partecipative, e sulla c.d. non esclusività della decisione algoritmica”*.

(3) Su alcuni aspetti critici dell'uso dell'IA nell'amministrazione della giustizia si veda V. Tenore, *I robot in giudizio! considerazioni sull'utilizzo di intelligenza artificiale (A.I.) da parte del magistrato: abdicare al ragionamento ed alla riflessione quotidiana a favore di un robot togato, ovvero un 'cretino digitale', è un bene per la giustizia e per i suoi attori e destinatari?*, in *Rivista Corte conti*, 2023, 4, 1.

(4) T. Tessaro, *I controlli finanziari sull'attività degli Enti territoriali tra modelli concreti di verifica della Corte dei conti ed effettiva capacità predittiva*, intervento svolto al convegno organizzato a Varese da Upel il 18 aprile 2024 nel corso del quale è stata presentata un'esperienza applicativa di IA presso la Sez. contr. reg. Emilia-Romagna.

(5) Art. 12, c. 1, lett. b), d.l. 16 luglio 2020, n. 76, convertito con modificazioni, dalla l. 11 settembre 2020, n. 120.

(6) Sui rischi e sulle sfide dell'istruttoria/decisione amministrativa algoritmica si è soffermata D.U. Galetta, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità. Rischi e sfide della rivoluzione tecnologica in atto*, in *www.federalismi.it*, 6 febbraio 2019.

(7) Sul tema si veda G.M. Racca, *La 'fiducia digitale' nei contratti pubblici tra piattaforme e data analysis*, in *Istituzioni del federalismo*, 2, 2023.

(8) Si veda, *ex multis*, Cons. Stato, Sez. VI, 4 febbraio 2020, n. 881.

(9) V. Neri, *Diritto amministrativo e intelligenza artificiale: un amore possibile*, in *Urb. e appalti*, 2021, 5, 581.

Ne deriva che, come sopra rilevato, la decisione adottata con ricorso all'algoritmo vede sempre la necessità che sia l'amministrazione a compiere un ruolo ex ante di selezione e verifica, anche per mezzo di costanti test, aggiornamenti e modalità di perfezionamento dell'algoritmo".

Può essere utile al riguardo richiamare la fattispecie oggetto dei giudizi definiti con tre sentenze gemelle del Consiglio di Stato (n. 8472, 8473 e 8474/2019) in cui il Ministero dell'Istruzione ha utilizzato una procedura automatizzata algoritmica nell'assegnazione delle sedi di servizio dei docenti in mobilità o assunti con reclutamento straordinario.

Il Consiglio di Stato in particolare ha censurato l'operato del ministero sia per la non conoscibilità dell'algoritmo o dell'atto amministrativo informatico *"con riferimento ai suoi autori, al procedimento utilizzato per la sua elaborazione comprensivo delle priorità assegnate nella procedura valutativa*, sia per gli esiti contraddittori ed illogici della stessa procedura come il trattamento di maggior favore riservato a docenti con minori titoli e minore anzianità.

4. IA ed i presupposti per configurare una responsabilità amministrativa-contabile

4.1. Ipotesi di danno erariale. C.d. indiretto

Da queste premesse possiamo quindi sempre più ipotizzare la lesione di diritti soggettivi e/o di interessi legittimi per effetto del non corretto utilizzo delle procedure automatizzate nell'adozione di 'decisioni algoritmiche' o nell'erogazione di servizi pubblici e, quindi, la soccombenza delle amministrazioni pubbliche nei giudizi risarcitori intentati dinanzi al giudice ordinario o al giudice amministrativo.

La soccombenza dell'amministrazione pubblica nei suddetti contenziosi determina, come sappiamo, il c.d. *danno erariale indiretto* che legittima, ex art. 28 della Carta costituzionale, l'azione di rivalsa del P.M. contabile nei confronti di coloro che hanno concorso con dolo o colpa grave nella causazione del danno erariale (10).

Tuttavia, se in astratto è possibile imputare una responsabilità per danno erariale conseguente al non corretto utilizzo di procedure automatizzate più o meno avanzate (*machine learning* o *deep learning*) certamente è assai complesso verificare i presupposti per imputare e comprovare una responsabilità amministrativa degli operatori pubblici in tale ambito.

In verità particolari problemi non si ritiene sussistano laddove il danno erariale sia conseguente a *malpractice* sanitaria, ad esempio per errata diagnosi, dove il medico non risponde ovviamente del malfunzionamento di un sistema di IA, ma solo del danno causato da un suo negligente e scorretto uso del sistema.

Il P.M. contabile in questi casi è chiamato a comprovare che il personale sanitario non si sia attenuto alle linee guida o alle buone pratiche clinico-assistenziali di cui all'art. 5 della legge Bianco-Gelli ed in dottrina è stato proposto di guardare alla indicazione algoritmica di un trattamento sanitario proprio come ad una linea guida medica (11).

4.2. IA e danno erariale diretto da attività amministrativa/provedimentale illegittima

Le maggiori criticità nell'accertamento della responsabilità amministrativa sono riscontrabili invece con riferimento ai danni erariali conseguenti all'utilizzo dell'IA nell'esercizio dell'attività amministrativa in senso stretto anche se, come è stato evidenziato in dottrina e nel recente Regolamento Ue n. 2024/1689, rimane sempre una 'riserva di umanità' nella scelta discrezionale della pubblica amministrazione, in virtù del principio di non esclusività del metodo algoritmico (12).

Per ovviare alle "scatole nere" che icasticamente governerebbero le decisioni artificiali il Regolamento in esame introduce l'obbligo di sorveglianza umana, *"human in the loop"*, cui consegue – si ribadisce – il principio di non esclusività della decisione algoritmica.

Tali criticità si rinvergono per l'opacità tecnica del *software* o del metodo algoritmico (si parla di *black box*) nelle ipotesi di IA generativa, dove è possibile conoscere i dati immessi (*input*) ed il risultato finale (*output*) ma non sempre è possibile comprendere il percorso 'logico' seguito dal *software* che ha portato alla determinazione conclusiva della macchina/calcolatrice.

Difficoltà sono riscontrabili anche nell'individuazione della persona a cui imputare la responsabilità amministrativa-contabile; si pensi alle ipotesi di istruttoria di un procedimento amministrativo incompleta o viziata. Sul punto è stato affermato che *"Ove l'istruttoria nella quale si esprime l'IA sia affidata a responsabile che non è legittimato ad adottare il provvedimento finale, si può configurare una lettura adattata dell'art. 6, lett. e) della legge n. 241 del 1990, con aggravamento dell'onere motivazionale del dirigente che intenderà discostarsi dalle risultanze dell'istruttoria. Tale facoltà, come noto immaginata per disgiungere i profili di responsabilità discendenti da istruttoria viziata o incompleta dalla produzione di effetti giuridici verso terzi del provvedimento, risulta ancor più complessa da esercitare ove l'istruttoria riposi su un sistema di IA. Il responsabile del procedimento predispone in tali casi uno schema di provvedimento costruito in ragione dell'output algoritmico, del quale dovrà aver controllato il corretto esercizio; il dirigente per potersi discostare dovrà affrontare un onere*

(10) N.M. Tritto, *Il danno erariale da intelligenza artificiale*, cit. p. 296 ss.; l'autore si sofferma su un'ampia casistica di 'danni erariali da IA', diretti ed indiretti.

(11) A.G. Grasso, *Diagnosi algoritmica errata e responsabilità medica*, in *Rivista di diritto civile*, 2023, n. 2, pp. 341 ss.; si veda anche G. Natale e F. D'Orazio, *La responsabilità medica alla prova dell'A.I.*, in *Rassegna Avvocatura dello Stato*, n. 1 2023, pp. 191 ss. Sul tema della responsabilità amministrativa del personale sanitario in generale si veda P. Evangelista – F. Foggia, *La responsabilità del medico davanti alla Corte dei conti dopo la legge n. 24/2017*, in *Monitor, Rivista di AGE.NA.S.*, 2023, n. 48.

(12) G. Gallone, *Riserva di umanità e funzioni amministrative. Indagine sui limiti dell'automazione decisionale tra procedimento e processo*, Padova, 2023.

motivazionale non solo limitato al valutato difetto istruttorio, ma esteso a vizi di funzionamento, lato sensu intesi, del sistema di IA” (13).

Altra criticità è rinvenibile nell’assenza di adeguate competenze e figure professionali ‘tecniche’ nelle piante organiche delle pubbliche amministrazioni, per cui non è semplice né corretto, pur richiamando la c.d. colpa professionale ex art. 1176, c. 2, c.c., imputare una condotta gravemente negligente al dipendente pubblico che non è in possesso di conoscenze per intercettare l’errore e l’illegittimità della ‘decisione algoritmica’.

Ed allora possiamo più propriamente contestare una responsabilità amministrativa contabile nelle ipotesi di gravi negligenze nelle fasi endoprocedimentali, propedeutiche alla decisione amministrativa automatizzata, tra cui la selezione e approvazione del software, dove è rilevante l’intervento ‘umano’ del dipendente pubblico che peraltro in tale ambito non può agire in violazione sia di norme interne che sovranazionali (Regolamento approvato dal Parlamento Europeo il 14 giugno 2024).

Si pensi al rispetto della disciplina normativa in materia di riservatezza dei dati acquisiti nel sistema (14) nonché al rispetto dei principi dettati dal richiamato art. 30 del nuovo codice dei contratti nell’adozione di decisioni mediante automazione quali:

- 1) conoscibilità e comprensibilità delle logiche di funzionamento del sistema;
- 2) non discriminazione algoritmica nei confronti di persone fisiche e degli operatori economici;
- 3) non esclusività della decisione algoritmica.

5. Conclusioni

Dalle premesse considerazioni e tenendo conto, come abbiamo visto, che deve essere sempre assicurato un controllo umano del procedimento algoritmico in funzione di garanzia (*il c.d. human in the loop*) possiamo pervenire alla conclusione che è possibile imputare ad un funzionario pubblico una responsabilità amministrativo-contabile per il danno erariale – diretto o indiretto – conseguente ad un non corretto utilizzo dell’IA, dovendo verificare a monte l’esattezza dei dati da elaborare e mantenere il costante controllo del procedimento ‘algoritmico’.

Parimenti gli amministratori pubblici non potrebbero invocare la c.d. “esimente politica” (15) nei casi di utilizzo dell’IA nella programmazione della spesa, nell’analisi predittiva dei dati di bilancio.

Sul punto la giurisprudenza contabile ha più volte ribadito che l’esimente non è applicabile nelle materie riservate agli organi di governo, nelle quali gli uffici amministrativi e tecnici della struttura abbiano espletato funzioni istruttorie o consultive e comunque di mero supporto strumentale (16).

È stato al riguardo precisato che “la *ratio* dell’esimente in questione risiede nella tutela della buona fede e dell’affidamento degli organi di governo in sede di adozione di atti deliberativi che necessitano di un’istruttoria tecnico-amministrativa particolarmente complessa, e consegue al principio di netta separazione e divieto di interferenza tra funzioni di indirizzo politico spettanti agli amministratori pubblici e attività gestorie di competenza dirigenziale” (17).

È necessaria, tuttavia, da parte del P.M. contabile e/o dall’Organo giudicante, una valutazione *cum grano salis* dell’elemento psicologico ovvero della condotta di chi ha concorso nell’emanazione dell’atto informatico illegittimo, fonte di danno erariale.

In assenza di strutture tecniche nella P.A. e di adeguate iniziative formative del personale amministrativo coinvolto nell’utilizzo di strumenti tecnologici, sempre più innovativi, appare indubbiamente difficile contestare una condotta gravemente negligente ed inescusabile dei dipendenti pubblici a fondamento della responsabilità amministrativo-contabile.

* * *

(13) S. Foa, *Intelligenza artificiale e cultura della trasparenza amministrativa. Dalle ‘scatole nere’ alla ‘casa di vetro’?*, in *Dir. amm.* 2023, 3, 515 ss.

(14) Per un commento sul provvedimento n. 112/2023 del Garante per la protezione dei dati personali si veda E. Maggio, *L’informativa per il trattamento dei dati personali di sistemi di intelligenza artificiale: il caso ChatGPT*, in *Nuovo diritto civile*, 2023, 4.

(15) L’art. 1 della l. 14 gennaio 1994, n. 20, disposizioni in materia di giurisdizione e controllo della Corte dei conti, come modificata dal d.l. 23 ottobre 1996, n. 543, convertito con modificazioni dalla l. 20 dicembre 1996, n. 639, prevede, al c. 1-ter, dispone che “[...] nel caso di atti che rientrano nella competenza propria degli uffici tecnici o amministrativi la responsabilità non si estende ai titolari degli organi politici che in buona fede li abbiano approvati ovvero ne abbiano autorizzato o consentito l’esecuzione”.

(16) *Ex multis*, si veda Corte conti, Sez. II centr. app., n. 111/2018; id. Sez. III centr. app. n. 129/2017). In Corte conti, Sez. giur. reg. Campania n. 314/2019 è stato affermato che “per pacifica giurisprudenza, la suddetta esimente non è applicabile nel caso di atti che, come l’approvazione del rendiconto di esercizio, sono assegnati dalla legge alla competenza di un organo politico” e “la presenza di pareri favorevoli e/o l’assenza di rilievi di legittimità da parte del Segretario-Direttore generale o dei Revisori dei conti non sono sufficienti a dispensare i componenti dell’organo politico dalle loro eventuali responsabilità” e che, in definitiva, “dispensare i componenti dell’organo politico da ogni responsabilità significherebbe praticamente ridurre, contrariamente alla volontà del legislatore, la deliberazione di Consiglio comunale di approvazione del rendiconto a mera ratifica di decisioni assunte da altri”.

(17) E. Tomassini, *Intelligenza artificiale, esimente politica e programmazione della spesa*, intervento nel corso del convegno organizzato a Varese nei giorni 18-19 aprile 2024 da Upel - Unione Provinciale Enti Locali, sul tema: intelligenza artificiale e appalti pubblici tra capacità predittiva e discrezionalità amministrativa.

INTELLIGENZA ARTIFICIALE E PUBBLICA AMMINISTRAZIONE

di Pierfrancesco Miele (*)

Abstract: Il contributo esamina i rischi sociali, economici e giuridici nascenti dal crescente utilizzo dell'intelligenza artificiale. Cos'è l'intelligenza artificiale e cosa si intende con tale locuzione, come funziona e quali sono le applicazioni che le persone, le imprese, le amministrazioni pubbliche, e, più in generale, la società ne possono fare, quali sono i vantaggi che ne possono trarre e quali sono i rischi che una indiscriminata applicazione della stessa può determinare per l'uomo, tra *machine learning*, *deep learning* e reti neurali. In particolare, esamina l'uso dell'intelligenza artificiale nel pubblico impiego per la valutazione del personale amministrativo, il conferimento di incarichi dirigenziali, nonché la verifica del conseguimento degli obiettivi e la valutazione delle *performances* dei dirigenti. Infine, sono illustrate la regolamentazione a livello europeo e la regolazione nel nostro Paese, nonché le linee guida del disegno di legge sull'intelligenza artificiale approvato dal Consiglio dei ministri il 23 aprile 2024.

The paper examines the social, economic and legal risks arising from the growing use of artificial intelligence. What is artificial intelligence and what is meant by this term, how does it work and what are the applications that people, businesses, public administrations and, more generally, society can make of it, what are the advantages that can draw from it and what are the risks that an indiscriminate application of the same can cause for humans, between machine learning, deep learning and neural networks. In particular, it examines the use of artificial intelligence in the public sector for the evaluation of administrative staff, the assignment of managerial positions, as well as the verification of the achievement of objectives and the evaluation of the performances of managers. Finally, the regulation at European level and the regulation in our country are illustrated, as well as the guidelines of the bill on artificial intelligence approved by the Council of Ministers on 23 April 2024.

Sommario: 1. Premessa. – 2. I rischi sociali, economici e giuridici nascenti dal crescente utilizzo dell'intelligenza artificiale. La relatività e la fallibilità dell'algoritmo e dell'intelligenza artificiale. – 3. Lavoro, occupazione e intelligenza artificiale. – 4. L'intelligenza artificiale nel campo del diritto. – 5. La riserva di umanità e la necessità di non abdicare alla primazia dell'intelligenza umana. – 6. Pubblica amministrazione e intelligenza artificiale. – 7. Intelligenza artificiale e attività amministrativa: la tutela dei diritti. – 8. Intelligenza artificiale, Big Data, e tutela della riservatezza dei dati personali (tutela della privacy). – 9. Intelligenza artificiale e pubblico impiego. – 10. Intelligenza artificiale, valutazione del personale amministrativo, conferimento di incarichi dirigenziali, conseguimento degli obiettivi e valutazione delle *performances* dei dirigenti e attribuzione del trattamento economico. – 11. Verso la regolazione dell'intelligenza artificiale: la regolamentazione a livello europeo e la regolazione nel nostro Paese. – 12. La regolamentazione a livello europeo dell'intelligenza artificiale: il Regolamento (UE) 2024/1689 del Parlamento europeo. – 13. La regolamentazione a livello nazionale dell'intelligenza artificiale: il disegno di legge recante "Disposizioni e delega al Governo in materia di intelligenza artificiale", approvato dal Consiglio dei ministri del 23 aprile 2024 e presentato al Senato il 20 maggio 2024 (A.S. 1146). – 14. Le linee guida del disegno di legge sull'intelligenza artificiale approvato dal Consiglio dei ministri il 23 aprile 2024. – 15. Il contenuto del disegno di legge sull'intelligenza artificiale approvato dal Consiglio dei ministri il 23 aprile 2024. – 16. Il grande limite della regolazione sull'utilizzo dell'intelligenza artificiale in ambito europeo e nazionale. Riflessioni conclusive e prospettive evolutive.

1. Premessa

Negli ultimi tempi si parla sempre più spesso di intelligenza artificiale, che possiamo ormai ritenere essere entrata nella nostra quotidianità. Di qui l'esigenza di chiederci che cos'è l'intelligenza artificiale e cosa si intende con tale locuzione, come funziona e quali sono le applicazioni che le persone, le imprese le amministrazioni pubbliche, e, più in generale, la società ne possono fare, quali sono i vantaggi che ne possono trarre e quali sono i rischi che una indiscriminata applicazione dell'intelligenza artificiale può determinare per l'uomo, tra *machine learning*, *deep learning* e reti neurali (1).

(*) P. Miele è Avvocato e Dottorando di ricerca in Diritto amministrativo presso l'Università degli Studi di Cassino e del Lazio meridionale.

(1) Sull'intelligenza artificiale esiste una vasta bibliografia riferita ai settori più disparati. Fra le opere più significative che hanno affrontato il tema dell'intelligenza artificiale in campo giuridico e al settore del diritto amministrativo in particolare si possono ricordare: M. Protto, *Diritto amministrativo e nuove tecnologie*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti* (a cura di), in *Giurisprudenza Italiana* - Giugno 2022, 1507 ss.; G. Pesce, *Diritto amministrativo e*

È sotto gli occhi di tutti che ormai l'intelligenza artificiale fa già parte della vita di moltissime persone in tutto il mondo e molteplici sono le applicazioni che di essa se ne fa in tutti i settori. In questo momento storico, infatti, sono diverse le tecnologie, gli strumenti e i servizi che ricorrono a una o più branche dell'*Artificial Intelligence*: strumenti e servizi che costruiti sul *machine learning*, sul *deep learning* e/o sulle reti neurali. Il risultato è che tantissime persone sfruttano l'intelligenza artificiale quotidianamente, anche senza rendersene conto; anche se magari, di fronte a una richiesta esplicita, non saprebbero spiegare che cos'è l'intelligenza artificiale. Proprio per questo è importante iniziare a conoscere in maniera più approfondita il funzionamento e le applicazioni di uno dei fenomeni scientifici più importanti del momento storico che stiamo vivendo.

L'intelligenza artificiale si presta infatti a utilizzi legati ai settori più disparati: si va dalla medicina, all'economia, al settore bancario, al campo giuridico, dall'attività amministrativa all'attività giudiziaria, con inevitabili problemi in termini di sicurezza dei dati personali delle persone interessate. Il tutto senza dimenticare che il fenomeno dell'intelligenza artificiale continuerà a crescere, aumentando i suoi ambiti d'uso e, di conseguenza, caratterizzando sempre di più il futuro di tutta la società in ogni settore della vita sociale. Un futuro all'insegna di nuove incredibili possibilità, ma anche di alcuni rischi in cui si potrebbe incappare, specie nel caso in cui l'intelligenza artificiale si continuasse a sviluppare senza adeguate forme di controllo.

Negli ultimi anni anche il settore pubblico, le istituzioni pubbliche e le pubbliche amministrazioni hanno mostrato una crescente attenzione verso i sistemi di intelligenza artificiale, promuovendo iniziative di ricerca e di sviluppo e sperimentandone l'impiego nel contesto della propria azione pubblica.

I campi applicativi già verificati sono molteplici e spaziano dalla gestione dei confini, all'attività di regolazione economica o al prelievo fiscale sino all'erogazione dei servizi pubblici e all'adozione di decisioni autoritative. Oltre a comportare la dematerializzazione dell'attività pubblica dei rapporti tra amministrazione e cittadini, tali processi sono suscettibili di modificare lo stesso potere amministrativo, nonché i modi del relativo esercizio.

Per cogliere l'importanza dell'intelligenza artificiale e il prevedibile sviluppo che la sua applicazione avrà in tutti i settori della vita sociale giova considerare i vantaggi che essa può comportare per le aziende e per le pubbliche amministrazioni. Non v'è dubbio, infatti che l'intelligenza artificiale aggiunge valore alle attuali capacità dei computer, fornendo in modo continuo e affidabile attività e informazioni computerizzate, aiutando a migliorare i risultati attraverso: a) processi e attività automatizzati; b) riduzione degli errori umani; c) incremento della produttività e dell'efficienza operativa; d) miglioramento delle decisioni operative con accesso a dati in tempo reale; e) migliore apprendimento dei dati con accesso potenziato a grandi quantità di dati; f) miglioramento del servizio grazie alla conoscenza del cliente o dell'utente; g) generazione di *lead* di qualità.

Entrando nello specifico, esistono cinque modi in cui l'intelligenza artificiale può essere utile per un'organizzazione: a) raccolta e analisi dei dati, essendo indubbio che l'intelligenza artificiale rende la raccolta e l'analisi dei dati economica, intuitiva e tempestiva, consentendo di acquisire automaticamente maggiori informazioni sui clienti e di assicurarsi business sia nuovo che consolidato; b) assunzioni più intelligenti, atteso che gli algoritmi di *machine learning* possono determinare le pratiche migliori per esigenze specifiche di assunzione e creare una short list dei candidati migliori; c) efficienza del *back office*, in quanto l'intelligenza artificiale può gestire attività quali contabilità, schedulazione e altre funzioni quotidiane in un attimo, senza errori; d) servizio clienti, o utenti, in quanto gli assistenti virtuali per il servizio ai clienti in una azienda, o il personale virtuale di una amministrazione pubblica lavorano 24 ore al giorno, 7 giorni alla settimana e possono aiutare clienti o utenti attivi e potenziali senza la supervisione di addetti umani; e) marketing mirato, potendo l'intelligenza artificiale consultare e classificare tutti i dati disponibili su un prodotto o servizio, consentendo all'azienda, o alla pubblica amministrazione di focalizzarsi su un *marketing* o su un servizio che individua specificamente le esigenze dei clienti o degli utenti.

Anche nel settore del pubblico impiego l'intelligenza artificiale può trovare, quindi, ampi spazi e ambiti di applicazione. Basti pensare ai vantaggi che l'uso dell'intelligenza artificiale può comportare nel controllo della spesa pubblica e nel controllo dei limiti di finanza pubblica, nella selezione del personale, soprattutto in termini

intelligenza artificiale: i problemi, in M. Protto, *Pubblica amministrazione e tecnologie emergenti* (a cura di), op. cit., 1508 ss.; D. Marangiu, *Algoritmo e procedimento amministrativo: una ricostruzione*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti* (a cura di), op. cit., 1515 ss.; A.G. Orofino, F. Cimbali, *L'uso delle tecniche informatiche nella prestazione di servizi pubblici*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti* (a cura di), op. cit., 1523 ss.; F. Costantino, *Algoritmi, intelligenza artificiale e giudice amministrativo*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti* (a cura di), op. cit., 1527 ss. Sempre sull'intelligenza artificiale in generale si veda ancora: M. Macchia, A. Mascio, *Intelligenza artificiale e sfera pubblica: lo stato dell'arte*, in *Giornale di diritto amministrativo*, 2022, 4, 556 ss.; M. Luciani, *La decisione giudiziaria robotica*, in *Nuovo diritto civile*, 2018, n. 1; V. Tenore, *I robot in giudizi. Considerazioni sull'utilizzo dell'intelligenza artificiale (AI) da parte del magistrato: abdicare al ragionamento e alla riflessione quotidiana a favore di un robot togato, ovvero un "cretino digitale", è un bene per la giustizia e per i suoi attori e destinatari*, in questa Rivista, 2023, 4, 1 ss.; M. Peruzzi, *Intelligenza artificiale, poteri datoriali e tutela del lavoro: ragionando di tecniche di trasparenza e poli regolativi*, in *Ianus, Diritto e Finanza, Rivista di Studi giuridici*, n. 24, Dicembre 2021, 71 ss. Sulla riserva di umanità si veda, in particolare: G. Gallone *Riserva di umanità e funzioni amministrative – Indagine sui limiti dell'automazione decisionale tra procedimento e processo*, Padova, Cedam, 2023; nonché G. Gallone *Digitalizzazione, amministrazione e persona: per una "riserva di umanità" tra spunti codicistici di teoria giuridica dell'automazione*, in *PA Persona e Amministrazione – Ricerche giuridiche sull'Amministrazione e l'Economia*, 329 ss.

di celerità dei processi di selezione e di efficienza nella valutazione dei candidati e nella scelta dei vincitori, nella selezione del personale in relazione ai diversi posti di funzione sulla base dei curricula e delle particolari attitudini di ciascuna unità di personale, nelle procedure di inquadramento e nelle progressioni. Si pensi ancora ai vantaggi che l'intelligenza artificiale può comportare nella valutazione del personale amministrativo o nella valutazione e nella verifica del conseguimento degli obiettivi e nella valutazione delle performances dei dirigenti, nelle procedure di mobilità, nel controllo del lavoro flessibile, o ancora ai vantaggi che l'intelligenza artificiale può comportare nella corretta ed equa attribuzione del trattamento economico, o nella selezione del personale nelle procedure di conferimento di incarichi dirigenziali, nella valutazione delle *performances* ai fini dell'attribuzione dei relativi incentivi economici. O si pensi, altresì, all'applicazione dell'intelligenza artificiale nella valutazione del rispetto del Codice di comportamento dei dipendenti pubblici, o nella istruttoria e nella fase decisione dei procedimenti disciplinari, o ancora all'applicazione dell'intelligenza artificiale nel controllo dei dipendenti pubblici che svolgano il c.d. lavoro agile.

Principalmente tre sono gli ordini di opportunità derivanti dall'utilizzo di dispositivi di IA nel settore pubblico: a) i guadagni in termini di efficienza interna alla pubblica amministrazione; b) il miglioramento dei processi decisionali; c) il rafforzamento dell'interazione tra cittadino governo, ivi compresi la fornitura di servizi migliori e più inclusivi ed il potenziamento dei meccanismi partecipativi dei cittadini alle attività del settore pubblico.

La considerazione di tali vantaggi non deve, però, portare a trascurare i rischi e gli aspetti critici che tutt'ora si collegano all'IA, specie se impiegata in ambito pubblico (errori; opacità dei processi algoritmici; dequotazione degli istituti di partecipazione). Come molti studi hanno già contribuito a dimostrare, l'efficienza che l'IA promette di assicurare potrebbe avere altrettante – e, in alcuni casi, più gravi – ripercussioni in termini di accountability, di trasparenza dei processi decisionali, nonché sulla stessa tutela dei diritti e delle libertà civili.

Alla luce di ambedue gli ordini di considerazioni, rifuggendo alla tentazione di una rigida opposizione tra tutela dei diritti ed efficienza, l'obiettivo deve essere quello di “governare con l'IA, governando l'IA”, con la finalità di assicurare lo sviluppo di sistemi di IA affidabili e operanti nel rispetto dei principi etici e dei diritti della persona.

2. I rischi sociali, economici e giuridici nascenti dal crescente utilizzo dell'intelligenza artificiale. La relatività e la fallibilità dell'algoritmo e dell'intelligenza artificiale

Con riferimento specifico ai rischi dell'intelligenza artificiale, molti studiosi ed analisti si interrogano da tempo su quali strumenti attivare per impedire che l'evoluzione della società verso un'economia a sempre minore intensità di lavoro – la cui evoluzione è oggi accelerata dall'intelligenza artificiale – non si traduca in un impoverimento della popolazione, situazione che richiederebbe una “redistribuire” della ricchezza considerando che la maggior parte di questa verrà prodotta dalle macchine.

Alle tematiche sociali, si affiancano, poi, questioni etiche sullo sviluppo e l'evoluzione dell'intelligenza artificiale e delle nuove tecnologie. Ci si interroga da tempo sul “potere degli algoritmi” e dei *Big Data*, domandandosi se questi segneranno la superiorità del cervello delle macchine su quello dell'uomo. I timori (alimentati in rete da noti personaggi di spicco come Stephen Hawking ed Elon Musk) possono apparire eccessivi ma sottovalutare gli impatti dell'intelligenza artificiale potrebbe rappresentare il rischio numero uno. A mettere in guardia dai rischi dell'intelligenza artificiale è stato, primo fra altri personaggi di spicco, il noto fisico Stephen Hawking, il quale, durante l'ultimo Web Summit di Lisbona ha avuto modo di affermare che “(...) *non siamo in grado di prevedere cosa riusciremo a fare quando le nostre menti saranno amplificate dall'intelligenza artificiale. Forse, con strumenti nuovi, riusciremo anche a rimediare a tutti i danni che stiamo provocando alla natura, e magari saremo anche in grado di trovare soluzioni definitive a povertà e malattie. Ma è anche possibile che con la distruzione di milioni di posti di lavoro venga distrutta la nostra economia e la nostra società. L'intelligenza artificiale potrebbe essere il peggior evento della storia della nostra civiltà* – è la visione drammatica dell'astrofisico, – *in quanto essa porta con sé pericoli, come potenti armi automatiche, nucleari o biologiche, addirittura abilita nuovi modi per permettere a pochi individui ed organizzazioni di opprimere e controllare moltitudini di uomini (e cose). Dobbiamo prepararci a gestirla per evitare che questi potenziali rischi prendano forma e diventino realtà*”.

3. Lavoro, occupazione e intelligenza artificiale

Quando si parla di intelligenza artificiale non si può non toccare, fra gli altri, anche aspetti etici e sociali come quelli legati al lavoro e all'occupazione, dato che i timori nella comunità globale crescono. Timori giustificati se si pensa che la metà delle attività lavorative di oggi potrebbe essere automatizzata entro il 2055.

Qualsiasi tipo di lavoro è soggetto a una automazione parziale ed è partendo da questa considerazione che nel report *A Future That Works: Automation, Employment and Productivity*, realizzato da McKinsey Global Institute-MGI (un report di 148 pagine, disponibile sul sito del *World Economic Forum* di Davos, dove è stato ufficialmente presentato nello scorso gennaio), si stima che circa la metà dell'attuale forza lavoro possa essere impattata dall'automazione grazie alle tecnologie già note e in uso oggi.

Sorprende anche che l'ultimo monito sia venuto, al riguardo, proprio da un imprenditore di successo come Elon Musk. "L'intelligenza artificiale è il più grande rischio cui la nostra civilizzazione si trova a far fronte", ha avvertito. In particolare, ha evidenziato i rischi di una guerra scatenata dai computer o una catastrofe occupazionale dovuta a decisioni basate soltanto sulle elaborazioni dell'intelligenza artificiale, unico vero pilastro dominante dell'economia del futuro capace di riservare alle macchine migliaia, forse milioni, di lavori oggi ancora gestiti agli uomini.

Come dimostra la realtà, l'algoritmo può fallire. Gli errori possono essere causati sia dagli inputs (i dati processati dal sistema) sia dai programmi del computer (la serie di algoritmi che trasformano gli *inputs* in *outputs*) (2).

La programmazione e l'organizzazione sono probabilmente la principale fonte di errori per le decisioni amministrative generate dai sistemi computerizzati. L'automatizzazione, infatti, richiede la traduzione di regole legali – per definizione generali – nel linguaggio. Comunque, anche quando le statistiche sono elaborate in maniera completa, quando fattori di differente complessità interagiscono tra di loro, ci sono sempre variabili che possono non essere previste (i cosiddetti "cigni neri").

Il tema della responsabilità derivante dall'agente artificiale assume interesse giuridico, anzitutto, perché a ben intendere si è al cospetto di una "intelligenza senza coscienza", per quanto in grado di operare in modo autonomo. La tesi dell'intelligenza "senza coscienza", che connota il robot, escluderebbe affascinanti paragoni con fattispecie di responsabilità per fatto attribuibile agli animali o ad altri soggetti che agiscono per conto del padrone (come lo schiavo dell'era romana). Ed è proprio nella verifica della permanente sussistenza del nesso causale tra condotta e evento che si annida il problema che, forse, non può essere risolto soltanto auspicando la maggiore trasparenza, conoscibilità e partecipazione del privato al procedimento di formazione della decisione pubblica e, probabilmente, neppure immaginando formule di preventiva limitazione del ruolo della tecnica.

Se l'algoritmo intelligente o super intelligente fornisce una risposta, grazie all'impiego della statistica, senza che a noi umani sia permesso di capire perché quella risposta è stata fornita (la sua *ratio*), non siamo in grado nemmeno di capire se e in quale passaggio si è verificato un errore. Sappiamo che il fulcro della responsabilità presuppone una condotta, un evento (dannoso) e il nesso di causalità. Al centro della riflessione sulla responsabilità, da sempre, viene collocata la volontà, per quanto sul piano giuridico la responsabilità subisce attenuazione a seconda del grado di coscienza del fatto causativo del danno da parte dell'agente. Nel verificare il grado di coscienza e volontà dell'evento stesso, la responsabilità dell'agente può anche attenuarsi sino a scomparire. Come è stato detto da alcuni analisti della materia, *"fino a qualche anno fa nessuno si sarebbe mai sognato di ritenere responsabile della morte di un uomo la sola automobile perché 'tecnicamente' è essa che ha causato l'evento morte urtando la persona; ebbene oggi, dinanzi ad auto senza guidatore o ad armi a guida automatica, una questione del genere può diventare molto più complessa"*.

Le norme esistenti in materia di responsabilità si rivelano quindi insufficienti, perché coprono i casi in cui la causa di un'azione o di un'omissione del robot può essere fatta risalire ad uno specifico agente umano (ad esempio il fabbricante, l'operatore, il proprietario o l'utilizzatore), laddove quest'ultimo avrebbe potuto prevedere ed evitare il comportamento nocivo del robot stesso.

È in atto una riflessione, che da noi può ancora essere ritenuta avveniristica, ma della quale è opportuno tenere conto proprio per comprendere la dimensione del fenomeno.

4. L'intelligenza artificiale nel campo del diritto

Come è stato osservato da autorevoli studiosi che per primi si sono dedicati all'applicazione dell'intelligenza artificiale al campo del diritto, *"l'intelligenza artificiale e gli algoritmi promettono neutralità, razionalità ed obiettività. Utilizzare ove possibile l'algoritmo permette di sostituire, con costi tutto sommato contenuti, la discrezionalità umana con un giudizio sintetico ritenuto, di norma, non contaminato dall'influenza del pregiudizio, le cui applicazioni sono vaste: dai contratti, al processo, all'organizzazione della macchina amministrativa, alle decisioni"* (3).

Il diritto privato, di conseguenza, si interroga sulle implicazioni legate all'impatto dell'IA sui contratti, sulla responsabilità per i danni connessi all'uso della tecnologia, sulla esistenza di altre entità artificiali cui estendere l'imputazione giuridica. Il diritto penale è chiamato, a sua volta, ad affrontare non solo i delicati profili della imputabilità ma anche l'impatto dell'IA sulla ricerca dei reati, sulla configurazione della personalità del reo, sui rapporti con la giurisdizione.

Il diritto costituzionale e la filosofia del diritto esaminano la potenziale dicotomia tra lo sviluppo di sistemi intelligenti e le libertà costituzionali, a partire dal grado di diffusione della connessione alla potenziale sostituzione delle tecnologie emergenti sulle libertà collettive e sui diritti sociali così concepiti nella nostra Costituzione,

(2) Cfr. G. Pesce *Diritto amministrativo e intelligenza artificiale: i problemi*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti* (a cura di), op. cit., 1513 ss.

(3) Cfr. G. Pesce *Diritto amministrativo e intelligenza artificiale: i problemi*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti* (a cura di), op. cit., 1508 ss.

per passare all'impatto sul principio di eguaglianza e, in tempi di pandemia, sul diritto alla salute. I processi tecnologici incidono così sui capisaldi del diritto e sulle sue fonti, sulla società e sul modo di intendere le relazioni umane.

Per ciò che riguarda il diritto amministrativo si pongono, a ben vedere, le sfide più importanti, in considerazione del fatto che la pubblica amministrazione annovera tra i suoi ruoli la soddisfazione delle scelte dei cittadini e di quelle collettive. A sua volta, il diritto amministrativo include tra le sue missioni la dotazione giuridica e gli strumenti necessari agli apparati pubblici per realizzare le funzioni assegnate dalle leggi.

5. La riserva di umanità e la necessità di non abdicare alla primazia dell'intelligenza umana

Fra gli studiosi che si sono fin qui occupati di intelligenza artificiale ci si è interrogati sul ruolo da riconoscere alla persona umana nelle procedure amministrative automatizzate e sulla esistenza di una "riserva di umanità" nello svolgimento delle funzioni amministrative segnalando come anche i più recenti interventi legislativi nel campo dell'evidenza pubblica (con il nuovo Codice dei contratti pubblici – d.lgs. n. 36/2023) depongano ormai in questa direzione (4).

Ci si è chiesto, in particolare, se nello svolgimento delle funzioni amministrative automatizzate da parte dei pubblici poteri esista una sfera giuridica incompressibile di appannaggio dell'essere umano e, quindi, perimetrarne portata e contenuto. La risposta – si è detto – *“non può che essere affermativa ed appare legata a quella che può definirsi, senza enfasi, un'opzione ordinamentale di fondo di carattere assiologico. Definire il ruolo della persona nel processo decisionale amministrativo equivale, infatti, prima di ogni ulteriore riflessione dogmatica, a confrontarsi ed interrogarsi, sul piano filosofico ed etico, in ordine al rapporto uomo-macchina, come declinazione del tema del rapporto tra soggetto ed oggetto io e non io”* (5).

In tale contesto si è auspicato, in particolare, l'affermazione di una sorta di un "umanesimo digitale", figlio della nostra epoca dell'"umanesimo giuridico", il cui principale corollario dovrebbe essere la necessità di ricondurre la macchina (e, prima ancora, l'algoritmo ed il software di cui esso è espressione) a mero *instrumentum* al servizio dell'uomo e l'assegnazione alla stessa di una funzione servente di supporto e non sostitutiva dell'umanità. Ciò si traduce nella ricerca di un punto di equilibrio tra la quantità di potere che si delega alla macchina e quello che si mantiene agli esseri umani. Esso è da individuare, secondo l'opinione più accreditata, attraverso un criterio di "meta-autonomia" che raccomanda che siano gli esseri umani ad avere il potere di decidere quali decisioni prendere, esercitando la libertà di scegliere, dove necessario, e cedendo tale libertà, mediante una scelta comunque revocabile, nei casi in cui ragioni di opportunità e convenienza possono ritenersi prevalenti.

Dal punto di vista tecnico la soluzione avanzata è quella della costruzione di un modello di intelligenza artificiale che sia *human-centered designed* e che preveda il c.d. *Human-in-the-loop*¹⁸ collocando la conoscenza e l'esperienza delle persone al centro dei processi di apprendimento automatico. In esso la rinnovata centralità dell'individuo va assicurata, in particolare, con riguardo al momento finale dell'assunzione della scelta, prevedendo che il risultato computazionale elaborato dal software, prima di rifluire in una decisione, possa essere sottoposto a revisione da parte di una persona fisica che sia a ciò preposta (6). In questo senso si è proposto di parlare, sul piano giuridico, di "riserva di umanità" (7) nello svolgimento delle funzioni amministrative con ciò indicando quella sfera minima di intervento che deve inderogabilmente rimanere di appannaggio esclusivo della persona (con il corollario del divieto di forme di automazione integrale).

(4) Sulla riserva di umanità nello svolgimento delle funzioni amministrative si veda, in particolare, G. Gallone *Riserva di umanità e funzioni amministrative – Indagine sui limiti dell'automazione decisionale tra procedimento e processo*, Padova, Cedam, 2023, cit., nonché G. Gallone *Digitalizzazione, amministrazione e persona: per una "riserva di umanità" tra spunti codicistici di teoria giuridica dell'automazione*, in *PA Persona e Amministrazione – Ricerche Giuridiche sull'Amministrazione e l'Economia*, 329 ss., cit. Sulla riserva di umanità nell'esercizio della funzione giurisdizionale si veda, in particolare, V. Tenore, *I robot in giudizio. Considerazioni sull'utilizzo dell'intelligenza artificiale (AI) da parte del magistrato: abdicare al ragionamento e alla riflessione quotidiana a favore di un robot togato, ovvero un "cretino digitale"*, è un bene per la giustizia e per i suoi attori e destinatari, in questa Rivista, 2023, 4, 1 ss., cit.

(5) Cfr. G. Gallone *Digitalizzazione, amministrazione e persona: per una "riserva di umanità" tra spunti codicistici di teoria giuridica dell'automazione*, cit., 332.

(6) Tale soluzione è stata proposta da L. Floridi, *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*, cit., 98, secondo cui *“gli esseri umani dovrebbero mantenere il potere di decidere quali decisioni prendere, esercitando la libertà di scelta dove necessario e cedendola nei casi in cui ragioni di primaria importanza, come l'efficacia, possano prevalere sulla perdita di controllo sul processo decisionale”* con l'ulteriore precisazione che *“qualsiasi delega dovrebbe anche rimanere in linea di principio rivedibile, adottando come ultima garanzia il potere di decidere di decidere di nuovo”*.

(7) L'espressione "riserva di umanità" (in spagnolo "reserva de humanidad") la si deve a J. Ponce Solè, *Inteligencia artificial, Derecho administrativo y reserva de humanidad algoritmos y procedimiento administrativo debido tecnológico*, in *Revista General de Derecho Administrativo*, 2019, 50. In Italia, riprendendo sul punto sempre J. Ponce Solè, hanno parlato di una "tradizionale riserva di umanità" propria del diritto F. Fracchia, M. Occhiena, *Le norme interne: potere, organizzazioni e ordinamenti. Spunti per definire un modello teorico-concettuale generale applicabile anche alle reti, ai social e all'intelligenza artificiale*, Napoli, Editoriale Scientifica, 2020, 137.

Questa scelta terminologica non si spiega solo per la forza evocativa della formula ma riposa su specifiche considerazioni dogmatiche. Anzitutto, il concetto di “riserva”, istituto classico del diritto pubblico transitato dalla tradizione liberale allo Stato costituzionale contemporaneo, esprime, nelle sue diverse declinazioni (di legge, di giurisdizione, di amministrazione), una precipua pregnanza giuridica individuando una sfera materiale o di azione non contendibile e di appannaggio esclusivo di uno specifico potere dello Stato ovvero di una specifica fonte del diritto. Inoltre, esso non è estraneo al campo amministrativo essendo stato impiegato ad altri fini in un’accezione che, tuttavia, in sostanza, è sempre quella di delimitazione di uno spazio giuridico minimo.

È stato altresì osservato come la locuzione “riserva di umanità” sembra, poi, restituire maggiore profondità rispetto ad altre espressioni, pur di successo e, come si dirà in chiusura, oggi penetrate nel tessuto normativo, quali la “non esclusività algoritmica”. Quest’ultima, in quanto costruita in termini puramente negativi come mero limite, non dà, infatti, contezza della portata positiva del concetto potendosi prestare ad una lettura oltremodo riduttiva che rischia di relegare l’intervento umano nel procedimento quasi a mera eccezione rispetto alla regola dell’informatizzazione-automazione.

In definitiva, gli studiosi che hanno, fin qui affrontato la tematica dell’intelligenza artificiale applicata all’esercizio dei pubblici poteri hanno evidenziato come la questione della c.d. riserva di umanità consista nello stabilire se, e in quali limiti, sia accettabile che le decisioni amministrative riguardanti noi cittadini possano essere prese sulla base di meccanismi di intelligenza artificiale – in altre parole, di algoritmi – senza l’intervento umano volto a verificare che tali decisioni non siano viziate, illegali, discriminatorie, ecc. (8).

Per delimitare correttamente la questione, – si è detto – sono necessarie tre precisazioni. La prima afferisce alla frequenza di decisioni amministrative integralmente automatiche, cioè decisioni amministrative che si limitano ad avallare la soluzione raggiunta da un algoritmo, senza alcun controllo umano finale. La verità è che, allo stato attuale, questo tipo di decisione è rara, o addirittura molto rara. Indubbiamente, però, ci sono casi in cui c’è una spiccata influenza dell’algoritmo nell’orientare le scelte delle amministrazioni, come accade per il calcolo delle imposte, che viene effettuato dai computer dell’amministrazione fiscale, ovvero la regolazione dei semafori in alcune città, che viene effettuata da un sistema centrale di controllo digitale del traffico o, ancora, la gestione delle applicazioni in alcune procedure di appalto pubblico automatizzate. Ma va aggiunto che, anche nelle ipotesi poc’anzi accennate, non è detto che l’intervento umano nella decisione finale sia del tutto escluso: quello dei funzionari del fisco se il calcolo delle imposte effettuato dal computer rivela dati molto anomali, quello degli agenti incaricati di monitorare il sistema del traffico, se le scelte effettuate dal computer portano a creare situazioni anomale e problematiche.

La seconda precisazione da fare riguarda il significato di “automazione” nel campo dell’informatica pubblica. Dobbiamo capire che non basta che un computer intervenga nell’elaborazione di una decisione pubblica perché possa parlarsi di vera automazione. O più precisamente: l’intervento del computer implica sempre un elemento di automazione, ma a volte riguarda calcoli molto semplici che i funzionari pubblici potrebbero fare da soli, ancorché con maggiore spendita di tempo ed energie. La questione dell’automazione va oltre: nell’intervento di algoritmi per forgiare una soluzione a un problema complesso che l’amministrazione deve risolvere, seguendo il metodo algoritmico che consiste nello scomporre questo problema in sequenze logiche, in modo da elaborare un codice sorgente che possa essere applicato ai dati a disposizione. È importante rendersi conto che l’intervento umano può avvenire, e di solito avviene, a diversi livelli: la decisione di utilizzare un algoritmo, lo sviluppo dell’algoritmo in dialogo con gli esperti, l’adozione dell’algoritmo e infine la decisione finale. Il problema della decisione “automatica” sorge nell’ultima fase, il che non significa che le negligenze e gli errori commessi dall’amministrazione nelle tre fasi precedenti siano irrilevanti. Tali errori ineriscono ad un differente profilo problematico, anche se la loro esistenza è suscettibile di aggravare la natura discutibile della decisione finale “automatica”.

La terza precisazione riguarda l’ambito delle decisioni amministrative che possono essere problematiche se prese in modo “automatico”. Si tende a concentrare l’attenzione sulle decisioni individuali, e ancora più specificamente su quelle che possono avere effetti negativi sui diritti dei cittadini. Tuttavia, è necessario prestare attenzione perché, da un lato, tutte le decisioni individuali sono potenzialmente in grado di produrre effetti sia sui loro destinatari che su terzi e, dall’altro, le decisioni di regolamentazione o di pianificazione sono spesso in grado di incidere sui diritti. Consideriamo, ad esempio, i vincoli che possono essere creati da un regolamento di polizia basato su una valutazione algoritmica del rischio: è accettabile che venga emesso senza che nessun funzionario pubblico o eletto metta in discussione ciò che l’algoritmo propone?

6. Pubblica amministrazione e intelligenza artificiale

Come già si è avuto modo di anticipare, l’algoritmo, la *blockchain* e l’intelligenza artificiale sono termini e strumenti con cui il giurista non può più evitare di confrontarsi. Per la scienza del diritto amministrativo l’approccio assume i caratteri dell’urgenza, poiché il campo elettivo dell’impiego e della sperimentazione delle tecnologie

(8) Si veda *Prefazione* di Jean-Bernard Auby, in G. Gallone *Riserva di umanità e funzioni amministrative – Indagine sui limiti dell’automazione decisionale tra procedimento e processo*, cit.

emergenti sono certamente i settori che vedono necessariamente coinvolta la pubblica amministrazione (9). Come per la globalizzazione anche il ricorso alle nuove tecnologie richiede una rivisitazione delle categorie tradizionali del diritto amministrativo (quelle del provvedimento e del procedimento amministrativo in prima battuta), ma soprattutto la necessità di conciliare l'utilizzo di tali tecnologie con le garanzie che il cittadino ha faticosamente conquistato nel rapporto con l'amministrazione.

Il giudice amministrativo ha recentemente ritenuto illegittima la scelta dell'Amministrazione di affidare la gestione della procedura di mobilità straordinaria per il personale docente ad un algoritmo, constatando l'assenza di una vera e propria attività amministrativa, *"(...) essendosi demandato ad un impersonale algoritmo lo svolgimento dell'intera procedura di assegnazione dei docenti alle sedi disponibili nell'organico dell'autonomia della scuola"*, e precisando che *"alcuna complicatezza o ampiezza, in termini di numero di soggetti coinvolti ed ambiti territoriali interessati, di una procedura amministrativa, può legittimare la sua devoluzione ad un meccanismo informatico o matematico del tutto impersonale e orfano di capacità valutazionali delle singole fattispecie concrete, tipiche invece della tradizionale e garantistica istruttoria procedimentale che deve informare l'attività amministrativa, specie ove sfociante in atti provvedimentali incisivi di posizioni giuridiche soggettive di soggetti privati e di conseguenziali ovvie ricadute anche sugli apparati e gli assetti della pubblica amministrazione"* (10).

Secondo i giudici, un algoritmo, anche se strutturato in modo da tener conto di posizioni personali, di titoli e punteggi, *"(...) giammai può assicurare la salvaguardia delle garanzie procedurali che gli artt. 2, 6, 7, 8, 9, 10 della legge 7 agosto 1990 n. 241 hanno apprestato, tra l'altro in recepimento di un inveterato percorso giurisprudenziale e dottrinario (...) gli istituti di partecipazione, di trasparenza e di accesso, in sintesi, di relazione del privato con i pubblici poteri non possono essere legittimamente mortificati e compressi soppiantando l'attività umana con quella impersonale, che poi non è attività, ossia prodotto delle azioni dell'uomo, che può essere svolta in applicazione di regole o procedure informatiche o matematiche"*. Ad essere inoltre vulnerato non è solo il canone di trasparenza e di partecipazione procedimentale, ma anche l'obbligo di motivazione delle decisioni amministrative *"con il risultato di una frustrazione anche delle correlate garanzie processuali che declinano sul versante del diritto di azione e difesa in giudizio di cui all'art. 24 Cost., diritto che risulta compromesso tutte le volte in cui l'assenza della motivazione non permette inizialmente all'interessato e successivamente, su impulso di questi, al giudice, di percepire l'iter logico – giuridico seguito dall'amministrazione per giungere ad un determinato approdo provvedimentale"* (11). In altri termini, *"l'utilizzo di procedure informatizzate non può essere motivo di elusione dei principi che conformano il nostro ordinamento e che regolano lo svolgersi dell'attività amministrativa. In tale contesto, infatti, il ricorso all'algoritmo va correttamente inquadrato in termini di modulo organizzativo, di strumento procedimentale ed istruttorio, soggetto alle verifiche tipiche di ogni procedimento amministrativo, il quale resta il modus operandi della scelta autoritativa, da svolgersi sulla scorta della legislazione attributiva del potere e delle finalità dalla stessa attribuite all'organo pubblico, titolare del potere"*. Sembra quindi che proprio quella visione relazionale, materiale ed umana, del rapporto amministrativo, frutto del lento ma inesorabile processo evolutivo della disciplina, si opponga all'automatismo e impersonalità del ricorso alle nuove tecnologie.

7. Intelligenza artificiale e attività amministrativa: la tutela dei diritti

Nell'esercizio delle funzioni amministrative e nello svolgimento dell'attività amministrativa mediante l'intelligenza artificiale, e con riferimento specifico al rapporto di pubblico impiego, vengono in rilievo profili attinenti alla tutela dei diritti del dipendente pubblico alla luce delle varie criticità che si sono segnalate nelle pagine precedenti.

Pur in considerazione degli indubbi vantaggi in termini di efficienza, completezza e celerità che l'automazione dei procedimenti amministrativi riguardanti il pubblico impiego può comportare e sicuramente comporterà, non devono però essere trascurati e devono essere tenuti nella dovuta considerazione gli aspetti critici che sono da collegare all'uso dell'intelligenza artificiale, specie se impiegata in ambito pubblico (bias; errori; opacità dei processi algoritmici; dequotazione degli istituti di partecipazione). Come molti studi hanno già contribuito a dimostrare, l'efficienza che l'IA promette di assicurare potrebbe avere altrettante – in alcuni casi, più gravi – ripercussioni in termini di accountability, di trasparenza dei processi decisionali nonché sullo stesso godimento dei diritti e delle libertà civili. Alla luce di ambedue gli ordini di considerazioni, rifuggendo alla tentazione di una rigida opposizione tra tutela dei diritti ed efficienza, l'obiettivo deve essere, quindi, quello di *"governare con l'IA, governando l'IA"*, con la finalità di assicurare lo sviluppo di sistemi di IA affidabili e operanti nel rispetto dei principi etici e dei diritti della persona.

(9) Cfr. G. Pesce, *Funzione amministrativa, intelligenza artificiale e Blockchain*, Napoli 2021.

(10) Si vedano: Tar Lazio, Roma, Sez. III-bis, 24 giugno 2021, n. 7589, che richiama le precedenti pronunce nn. 9224, 9225, 9226, 9227, 9228, 9229 e 9230/2018, nn. 4649, 4650, 5139, 10963 e 10964/2019, con conferma in appello ad opera delle pronunce nn. 8472/2019, 8473/2019, 8474/2019 e 881/2020).

(11) Cfr., da ultimo, Cons. Stato, Sez. VI, sent. n. 881/2020.

Tenuto conto dei molteplici aspetti critici, per orientare virtuosamente il funzionamento dell'IA nei molteplici aspetti riguardanti specificamente il rapporto di pubblico impiego, e declinarlo al perseguimento delle finalità di interesse pubblico nel pieno rispetto delle persone interessate, e per ciò che qui ne occupa, nel pieno rispetto delle tutele e dei diritti dei dipendenti pubblici, occorre una regolazione proattiva *ad hoc*, in grado di direzionare lo sviluppo dei processi in atto e correggere quelle distorsioni capaci di riflettersi sul godimento dei diritti e delle libertà individuali.

8. Intelligenza artificiale, Big Data, e tutela della riservatezza dei dati personali (tutela della privacy)

Uno dei problemi più importanti ed urgenti che le amministrazioni e gli enti pubblici devono affrontare nello svolgimento delle loro funzioni e nella gestione delle procedure e dei molteplici aspetti riguardanti specificamente il rapporto di pubblico impiego, è oggi costituito dall'esigenza di coniugare la corretta gestione della gran mole di dati personali dei cittadini, molti dei quali sensibili in quanto riguardanti aspetti sanitari, o fiscali, o previdenziali, o religiosi, o di identità sessuali, la loro gestione mediante sistemi automatizzati di intelligenza artificiale, e la tutela dei dati personali alla luce della normativa in materia di tutela della *privacy* prevista dal Regolamento 2016/679/UE (Gdpr) e dal d.lgs. n. 196/2003. La soluzione del problema è diventata urgente alla luce dell'esigenza di gestire in maniera corretta ed efficiente, oltre che sicura, i dati contenuti nelle numerose banche dati gestite dalla pubblica amministrazione, proteggendoli dall'attacco di hackers che hanno fatto della loro violazione e della estrapolazione dei dati in esse contenuti un nuovo ed assai diffuso settore di attività criminale.

Il fenomeno, invero, va analizzato con riferimento al processo decisionale, alla struttura organizzativa e alla *governance* delle amministrazioni e delle aziende pubbliche, nonché con riferimento alla gestione dei *Big Data* mediante sistemi di intelligenza artificiale e di semplificazione dei sistemi informativi delle pubbliche amministrazioni, oltre che con riferimento alle ricadute che il fenomeno inevitabilmente ha sulla tutela della riservatezza dei dati personali dei dipendenti pubblici.

Negli ultimi anni, l'esigenza di coniugare la corretta gestione della grande mole di dati personali dei dipendenti pubblici, molti dei quali sensibili, la loro gestione in maniera automatizzata, e quindi attraverso sistemi di intelligenza artificiale, e la tutela dei dati personali alla luce della normativa in materia di *privacy*, ha assunto nelle amministrazioni e nelle aziende pubbliche una priorità assoluta e di indifferibile soluzione. In una società e in amministrazioni pubbliche che hanno eletto a loro fondamento la conoscenza delle informazioni, assumono, pertanto, sempre più rilievo l'acquisizione, lo studio, l'analisi e la gestione dei dati. Questi processi, tuttavia, per la gran mole di dati da acquisire e gestire, hanno reso inadeguate le tradizionali modalità di analisi e richiedono una gestione mediante le nuove tecnologie *digital* che utilizzino sistemi automatizzati di intelligenza artificiale. Le nuove tecnologie soccorrono in questo senso, rendendo possibile il trattamento simultaneo di molti dati contemporaneamente e l'estrazione di informazioni da esse: è il fenomeno dei "*Big Data*". La centralità e le potenzialità a livello giuridico ed economico del fenomeno sono state prontamente percepite sia dagli operatori del diritto e dagli operatori economici, che dalle istituzioni. Secondo la Commissione Europea «i dati sono il fulcro dell'economia e della società della conoscenza del futuro». Una maggiore e migliore analisi dei dati consente di acquisire informazioni utili in vista di attività decisionali tanto per le istituzioni, per le amministrazioni pubbliche e i loro dipendenti, quanto per le aziende e per i cittadini. Da ciò deriva uno sforzo, anche economico, da parte delle amministrazioni e delle aziende pubbliche in questo ambito al fine di seguire processi decisionali e operare decisioni di diverso tipo, dalle politiche di business agli aspetti organizzativi e di gestione del personale. In considerazione di ciò occorre approfondire quest'ultimo profilo – l'uso dei *Big Data* nella gestione dei dati da parte delle amministrazioni e delle aziende pubbliche con sistemi di intelligenza artificiale – al fine di ricostruire l'impatto di queste pratiche, attualmente in via di diffusione, sulla efficienza amministrativa e sulla contestuale esigenza di tutelare la riservatezza dei dati personali trattati.

Sebbene le tecniche di intelligenza artificiale consentano di prendere decisioni migliori e in tempi più rapidi, è comunque importante sottolineare come esse presentino ancora problemi in termini di trasparenza e comprensibilità, di sicurezza e di tutela dei dati personali. La soluzione del problema è diventata urgente alla luce dell'esigenza di gestire in maniera corretta ed efficiente, oltre che sicura, i dati contenuti nelle numerose banche dati gestite dalla pubblica amministrazione, proteggendoli dall'attacco di *hackers* che hanno fatto della loro violazione e della estrapolazione dei dati in esse contenuti un nuovo ed assai diffuso settore di attività criminale. Il recente caso di hackeraggio del sistema informatico di varie strutture sanitarie e di alcune aziende sanitarie è, in tal senso, sicuramente emblematico.

9. Intelligenza artificiale e pubblico impiego

Alla luce di quanto si è fin qui detto, è indubbio che l'intelligenza artificiale già trova, e troverà sempre più in un prossimo futuro, ampi spazi nelle amministrazioni e negli enti pubblici nella gestione delle procedure e dei procedimenti riguardanti il rapporto di pubblico impiego con i propri dipendenti. In tale ambito, pur in considerazione degli indubbi vantaggi che l'uso degli algoritmi e dell'intelligenza artificiale potrà comportare nella gestione del rapporto di pubblico impiego dei propri dipendenti, occorrerà considerare i molteplici profili che l'uso

dell'intelligenza artificiale applicato al rapporto di pubblico impiego potrebbe porre in termini di trasparenza, di tutela dei diritti, di corretta valutazione, di comprensibilità, di sicurezza e di tutela dei dati personali.

Le forme di automazione rappresentano oggi un'esperienza comune e regolare nell'attività procedimentale e decisionale delle pubbliche amministrazioni nella gestione del rapporto di pubblico impiego dei propri dipendenti, con un uso trasversale che investe molteplici settori dell'azione degli enti pubblici: esempi (fra i molti possibili) sono le valutazioni automatizzate delle prove preselettive nei procedimenti concorsuali, le procedure di inquadramento e per le progressioni di carriera, la valutazione del personale amministrativo in costanza di rapporto di impiego per finalità varie, le procedure per il conferimento di incarichi dirigenziali o per la valutazione del conseguimento degli obiettivi e per la valutazione delle *performances* dei dirigenti e l'attribuzione del trattamento economico, o ancora le valutazioni nelle procedure di mobilità, o del lavoro flessibile e del lavoro agile, o, infine, le valutazioni del personale con riferimento al codice di comportamento e ai procedimenti disciplinari.

Pur nella loro eterogeneità, i procedimenti gestiti in modo automatizzato presentano problematiche comuni, la cui soluzione deve essere ricercata nell'impianto normativo vigente e nella giurisprudenza oggi più sedimentata e consolidata. In particolare, le questioni che devono essere affrontate in relazione all'elaborazione elettronica dei provvedimenti amministrativi sono quasi interamente incentrate su come debba essere adattata la disciplina del procedimento amministrativo all'ipotesi in cui il decisore pubblico agisce mediante algoritmi. Nello specifico, occorre ottenere elementi di certezza su quale sia la natura giuridica del software, chi sia preposto a controllarne l'azione, come si applichino ad un'istruttoria algoritmica le garanzie di partecipazione, accesso e trasparenza previste dalla legge sul procedimento, chi debba essere identificato quale responsabile dell'emanazione di un provvedimento emesso per il tramite della macchina. Nel momento in cui si indaga sulle soluzioni giuridiche a tali interrogativi, è possibile riscontrare un fatto peculiare e determinante nella ricostruzione complessiva, ovvero il ruolo preponderante della giurisprudenza, rispetto alla normativa, nell'inquadramento giuridico dei profili dell'automazione amministrativa. Dunque, in materia di procedimento automatizzato, alla mancanza di una normativa articolata ha sopperito, soprattutto negli ultimi anni, un ruolo incisivo della giurisprudenza amministrativa, la quale ha avuto modo di dedicare un numero crescente di sentenze a questioni relative all'amministrazione per algoritmi, in cui spesso ricorrono indicazioni e obiter dictum che vanno oltre i casi specifici, con una evidente funzione nomofilattica dei Tar e del Consiglio di Stato (12).

Un aspetto su cui la giurisprudenza si è maggiormente soffermata riguarda la gestione operativa del procedimento algoritmico, e nello specifico al ruolo di controllo e verifica dell'operatore umano sull'attività della macchina. Infatti, una giurisprudenza ricorrente, costante e convergente dei Tar e del Consiglio di Stato, ha individuato alcuni importanti criteri circa le modalità di conduzione di un procedimento basato su un algoritmo. Ciò è emerso in primo luogo nel contesto di una pluralità di sentenze relative alla formazione automatizzata delle graduatorie per l'assegnazione delle sedi agli insegnanti della scuola secondaria. Si tratta di un insieme di decisioni ampiamente noto ed esaminato a più riprese dalla dottrina, di cui qui interessa focalizzare un profilo: l'enunciazione, da parte del giudice amministrativo, di un principio che può essere definito di sussidiarietà dell'uomo rispetto alla macchina, in ragione del quale un procedimento può sì essere affidato ad un algoritmo, ma deve essere nel contempo costantemente presidiato da un funzionario umano, che supervisioni in concreto l'azione del computer ed eventualmente intervenga laddove la macchina presentasse dei limiti, per esempio nell'eventualità in cui l'algoritmo non contempli un caso specifico e non sia in grado di inquadralo nelle griglie dei propri parametri. La considerazione del giudice è riassunta in un passaggio della sentenza del Tar Lazio, Roma, n. 10963/2019 che a sua volta riprende quanto già affermato in decisioni precedenti, e precisamente che *"(...) le procedure informatiche, finanche ove pervengano al loro maggior grado di precisione e addirittura alla perfezione, non possano mai soppiantare, sostituendola davvero appieno, l'attività cognitiva, acquisitiva e di giudizio che solo un'istruttoria affidata ad un funzionario persona fisica è in grado di svolgere e che pertanto, al fine di assicurare l'osservanza degli istituti di partecipazione, di interlocuzione procedimentale, di acquisizione degli apporti collaborativi del privato e degli interessi coinvolti nel procedimento, deve seguire ad essere il dominus del procedimento stesso, all'uopo dominando le stesse procedure informatiche predisposte in funzione servente e alle quali va dunque riservato tutt'oggi un ruolo strumentale e meramente ausiliario in seno al procedimento amministrativo e giammai dominante o surrogatorio dell'attività dell'uomo"*.

Come si vede, pur non escludendo la possibilità del ricorso all'algoritmo nel procedimento amministrativo, e non volendola disincentivare in sé stessa, si pone però la condizione del presidio umano sulla sua azione, in modo che essa non si possa ritenere devoluta in modo esclusivo alla procedura automatizzata. Ciò, allo scopo ultimo di garantire i principi di partecipazione, di trasparenza e di collaboratività che mai devono venire meno.

È opportuno osservare che questa posizione del giudice non si riferisce solo alla possibilità che l'algoritmo presenti disfunzioni e dunque produca elementi di illegittimità; bensì ricomprende anche l'ipotesi in cui le istruzioni del programma informatico siano perfette, ovvero *"finanche ove pervengano al loro maggior grado di precisione"*.

(12) Cfr. D. Marangiu, *Algoritmo e procedimento amministrativo: una ricostruzione*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti (a cura di)*, op. cit., 1519.

In questa chiave, dunque, l'automazione configura, nel dispiegarsi del procedimento amministrativo, una dinamica del ruolo uomo-macchina che è di sussidiarietà e di complementarietà, la quale implica un ruolo rinnovato della figura del responsabile del procedimento il quale, nella conduzione dell'iter automatizzato, diviene il garante primario di questo argine ad una gestione spersonalizzata del procedimento stesso.

Una sentenza che conferma e ribadisce questo profilo è la decisione del Tar Lazio, Roma, n. 4409/2021, relativa al processo automatizzato di formazione delle graduatorie per l'accesso a corsi universitari a numero chiuso. Il giudice, richiamando anche in questo caso la propria giurisprudenza pregressa, sancisce che deve essere censurata: “(...) *la devoluzione in toto a siffatti meccanismi [gli algoritmi] quantunque in ipotesi perfetti sul piano delle scienze matematiche, del procedimento amministrativo nelle sue fasi cruciali involgenti precipuamente l'istruttoria, la motivazione, l'individuazione dei controinteressati, con abdicazione da parte del responsabile del procedimento del suo ruolo di timoniere ed anche di correttore, grazie allo svolgimento di supplementi di istruttoria e il disimpegno di interventi correttivi in riparazione di possibili distorsioni o aberrazioni, dell'esito provvedimentale conclusivo*”. Questo passaggio riprende l'idea per cui la necessità del presidio umano sull'azione della macchina sussiste anche nel caso in cui si ha la certezza del buon funzionamento dell'algoritmo (“quantunque in ipotesi perfetti sul piano delle scienze matematiche”), ma soprattutto, incentra il discorso sulla figura del responsabile del procedimento che non solo non vede il suo ruolo ridimensionato, ma anzi trova una piena collocazione in questa dinamica. Ciò avviene nello specifico con due funzioni, una colloquiale-relazionale, l'altra incisiva-correttiva.

Nel primo ruolo, relazionale, il responsabile ha il compito di “non disumanizzare” il procedimento e far sì che sia garantita e conservata la dimensione comunicativa con i cittadini interessati. Quindi il responsabile del procedimento deve comunque essere in ogni momento il referente degli interessati, e quindi fornire informazioni sullo stato di avanzamento della procedura, e soprattutto deve fornire ai privati qualunque indicazione circa il funzionamento dell'algoritmo, ovvero “tradurne” il contenuto in linguaggio intellegibile ed esplicitarne le dinamiche.

Nel secondo ruolo, sussidiario rispetto alla macchina, il responsabile del procedimento assume la funzione di supervisore dell'azione del *software* al fine di verificare qualora sia necessario lo svolgimento di supplementi di istruttoria ed effettuare interventi correttivi in riparazione di possibili distorsioni o aberrazioni, dell'esito provvedimentale conclusivo.

Dunque spetta al responsabile del procedimento *in primis* assicurare che non si verifichino circostanze in cui casi concreti non siano pienamente inquadrabili nelle griglie dell'algoritmo, e, in questi casi, è suo compito intervenire, disponendo se necessario la modifica dell'algoritmo stesso, al fine di consentire che le situazioni siano pienamente ricomprese, oppure predisponendo una decisione “non algoritmica” in sostituzione di un'azione che, se condotta dal *software*, difetterebbe di logica e quindi sarebbe viziata. La conclusione di questa fase, di “istruttoria algoritmica”, è seguita dal transito al momento decisivo finale (13).

10. *Intelligenza artificiale, valutazione del personale amministrativo, conferimento di incarichi dirigenziali, conseguimento degli obiettivi e valutazione delle performances dei dirigenti e attribuzione del trattamento economico*

Altre procedure in cui potrebbero trovare opportuna ed efficace applicazione gli algoritmi e l'intelligenza artificiale sono quelle relative alla valutazione del personale amministrativo, alla valutazione e alla selezione del personale per il conferimento di incarichi dirigenziali, alla valutazione del conseguimento degli obiettivi, alla valutazione delle *performances* dei dirigenti e alla attribuzione del relativo trattamento economico.

Nell'ambito del rapporto di pubblico impiego, in numerosi casi le amministrazioni pubbliche sono chiamate a valutare il personale dipendenti per finalità diverse, che possono andare dalla attribuzione di benefici economici alla formazione di graduatorie ai fini di un trasferimento o della assegnazione a determinati posti di funzione. Ebbene, il più delle volte e tradizionalmente la valutazione del personale per tali finalità è stata fatta in maniera empirica, con ampi margini di discrezionalità da parte della stessa amministrazione datoriale e di chi per essa è chiamato ad operare tali valutazioni, e con inevitabili errori di valutazione e conseguenti fasi di contestazione delle valutazioni medesime da parte del personale interessato in sede contenzioso, e con dispendio di tempo, di risorse finanziarie e di personale.

In tali casi, invero, l'impiego di algoritmi e dell'intelligenza artificiale non solo arrecherebbe indubbi vantaggi in termini di tempo, accelerando in maniera assai determinante le procedure di valutazione, ma riuscirebbe a garantire altresì quella certezza, quella trasparenza, quella equità di valutazione, quella correttezza, e quella legittimità che i metodi empirici tradizionali di selezione non riescono a garantire.

Allo stesso modo la valutazione e la selezione del personale per il conferimento di incarichi dirigenziali, come pure la valutazione del conseguimento degli obiettivi da parte degli stessi dirigenti, la valutazione delle *performances* e l'attribuzione del relativo trattamento economico, se operate attraverso il ricorso agli algoritmi e all'intelligenza artificiale ben potrebbe garantire indubbi vantaggi in termini di celerità, di efficienza, di legittimità e di corretta valutazione, con indubbi vantaggi anche in termini di riduzione del contenzioso e della relativa spesa che

(13) Cfr. D. Marangiu, *Algoritmo e procedimento amministrativo: una ricostruzione*, in M. Protto, *Pubblica amministrazione e tecnologie emergenti (a cura di)*, op. cit., 1520 ss.

assai spesso conseguono alle valutazioni operate in tali settori attraverso i tradizionali metodi di valutazione “umana” ed empirica.

11. *Verso la regolazione dell'intelligenza artificiale: la regolamentazione a livello europeo e la regolazione nel nostro Paese*

Come già si è detto nei capitoli precedenti, il crescente sviluppo delle applicazioni, in ogni settore della società e della vita di tutti i giorni, dei sistemi di intelligenza artificiale, e la consapevolezza dei rischi e dei pericoli che un utilizzo indiscriminato degli stessi comporta per la tutela dei diritti individuali delle persone, e segnatamente della tutela della riservatezza dei dati personali, ha reso sempre più urgente ed ormai non più differibile una regolamentazione dell'utilizzo dell'intelligenza artificiale, soprattutto in considerazione della sempre più diffusa applicazione della stessa alle banche dati e ai *Big Data* in esse presenti. Alla sempre maggiore diffusione dell'intelligenza artificiale in tutti i settori della vita sociale, dalla medicina alla comunicazione, alla sicurezza, alla previdenza, al mondo dello spettacolo, alla meteorologia, alla finanza e a tanti altri settori della vita sociale, ha corrisposto, peraltro, negli ultimi anni una costante e sempre maggiore applicazione dell'intelligenza artificiale anche a tutte le attività delle pubbliche amministrazioni. Basti pensare, infatti, alla normale gestione di tutte le procedure amministrative connesse allo svolgimento delle normali attività amministrative delle varie amministrazioni, o alla gestione delle attività connesse alla erogazione dei servizi pubblici.

La sempre maggiore diffusione dell'intelligenza artificiale nella pubblica amministrazione è stata accompagnata dalla consapevolezza che, se da un lato essa comporterà una maggiore efficienza, efficacia e velocizzazione dell'azione amministrativa, dall'altro comporta seri rischi per la tutela dei diritti fondamentali, e segnatamente, della tutela della riservatezza dei dati personali dei soggetti interessati, anche, e soprattutto, con riferimento alla possibile vulnerabilità delle banche dati utilizzate dalle stesse pubbliche amministrazioni, con il conseguente rischio di finire per costituire una vera e propria pericolosità sociale se si considera che l'aumento di capacità di calcolo, e di conseguente invasività, dei supercomputer, tanto che non si mancò di evidenziare i possibili rischi per la stessa democrazia.

Il rischio maggiore evidenziato ormai a livello mondiale è costituito dal fatto che l'uso indiscriminato e senza regole da valere a livello mondiale dei sistemi di intelligenza artificiale – e quindi dei sistemi di supercalcolo e *Big Data*, aventi capacità autonoma di apprendimento e di autogenerazione di soluzioni sociali – posti fuori da ogni forma di controllo etico e democratico, possano portare in tempi rapidi a soluzioni e quadri decisionali usabili contro il benessere dell'umanità. Il principale rischio connesso con l'intelligenza artificiale è quindi che queste tecnologie vengano utilizzate contro gli stessi cittadini, per disporre un controllo dei loro comportamenti, o ancor più preconstituendo false realtà con cui condizionare le loro decisioni individuali e collettive. Rischi che si riferiscono direttamente alla democrazia e alle condizioni di benessere e libertà in cui i cittadini svolgono la propria vita.

Queste preoccupazioni, e l'obiettivo esistenza reale dei rischi sopra descritti, hanno indotto negli ultimi anni a diverse formulazioni di norme, più o meno volontarie, di regolazione delle tecnologie di IA, a partire dai cosiddetti Principi di Asilomar (2017) fino al Codice Etico dell'Unione Europea (2019), confluito poi nell'*EU Artificial Intelligence Act* di recente approvazione. Tutto ciò con l'obiettivo di fare in modo che l'uso dell'intelligenza artificiale debba avere solidi riferimenti etici, che ponendo al centro il benessere delle persone e delle loro comunità eviti usi illegittimi di questo potente strumento per schedare, e quindi controllare, le persone e i loro comportamenti politici, professionali o personali o comunque per distorcerne capacità decisionali e libertà individuali e collettive, richiedendo trasparenza, affidabilità e imparzialità nella cumulatione dei dati e nella formulazione degli algoritmi e nell'uso di tali *Big Data System*, con garanzie di sicurezza, inclusione e privacy, garantite da responsabilità certe e verificabili nelle diverse fasi dell'intero processo.

Alla luce dei rischi sopra evidenziati, è maturata quindi la consapevolezza, seguita da iniziative concrete, di “governare con l'IA, governando l'IA”, con la finalità di assicurare lo sviluppo di sistemi di IA affidabili e operanti nel rispetto dei principi etici e dei diritti della persona (14).

Si è così assistito a varie iniziative di regolazione sia a livello europeo che a livello nazionale dell'intelligenza artificiale con l'obiettivo di orientare virtuosamente il funzionamento dell'IA e declinarla al perseguimento delle finalità di interesse pubblico ma attraverso una irrinunciabile ed ormai improcrastinabile regolazione dell'utilizzo dei sistemi di intelligenza artificiale.

Al riguardo giova ricordare che il primo passo per la regolazione dell'IA a livello europeo è rappresentato dall'approvazione, nel febbraio 2017, della Risoluzione “con raccomandazioni alla Commissione per le norme di diritto civile sulla robotica” da parte del Parlamento europeo. La Risoluzione invitava gli Stati membri a disciplinare in maniera omogenea gli aspetti civilistici della robotica e suggeriva la creazione di un'Agenzia europea per la robotica e l'IA composta da regolatori ed esperti esterni in grado di assicurare le competenze tecniche, etiche e normative necessarie per sostenere gli attori pubblici interessati, sia a livello dell'Unione che degli Stati membri,

(14) M. Macchia, A. Mascolo, *Intelligenza artificiale e sfera pubblica: lo stato dell'arte*, in *Giornale di diritto amministrativo*, 2022, 4, 556 ss.

nel loro sforzo di garantire una risposta tempestiva, etica e ben informata alle nuove opportunità e alle nuove sfide”. La Risoluzione auspicava una serie di cautele etiche, per assicurare che lo sviluppo e l'utilizzo dei robot avvenga in condizioni tali da preservare la dignità, l'autonomia e l'autodeterminazione degli individui, la tutela della privacy e che si presti attenzione “alla possibilità che nasca un attaccamento emotivo tra gli uomini e i robot, in particolare per i gruppi vulnerabili (bambini, anziani, disabili), per attenuare gli impatti emotivi e fisici. A distanza di un anno, con le due comunicazioni del 25 aprile 2018 e del 7 dicembre 2018, la Commissione europea ha delineato tre pilastri fondamentali su cui avrebbe dovuto basarsi la regolazione dell'IA (15).

A tali primi documenti programmatici ha fatto seguito la pubblicazione, nel mese di febbraio 2020, del “Libro bianco sull'intelligenza artificiale”, con lo scopo di fissare gli obiettivi ed i traguardi per la realizzazione di una società europea in cui soluzioni digitali, tecnologie affidabili, persone e imprese possano convivere in modo sinergico e sicuro, in un ambiente aperto, democratico, nell'ambito di un'economia dinamica e sostenibile. La posta in gioco della sfida – come già si è avuto modo di dire – sono diritti fondamentali, quali il diritto alla privacy, alla dignità umana, alla libertà di espressione e alla non discriminazione, ma anche il futuro economico dell'Europa. Per raggiungere questo obiettivo il Libro bianco individua tre direttrici principali, vale a dire: la necessità di un approccio sovranazionale, uno sforzo regolatorio proattivo e, in ultimo, un approccio *risk based*.

L'evoluzione del quadro normativo a livello europeo sull'intelligenza artificiale è esitata, il 13 marzo 2024, con l'approvazione in prima lettura, da parte del Parlamento europeo del Regolamento sull'intelligenza artificiale (IA), il c.d. *Artificial Intelligence Act* (AI Act), avente il fine di proteggere i diritti fondamentali, promuovendo l'innovazione, con particolare attenzione alle piccole e medie imprese, comprese le *start-up* e assicurando al contempo all'Europa un ruolo guida nell'adozione di un'IA affidabile ed etica.

La proposta di Regolamento era stata presentata dalla Commissione il 21 aprile 2021 come parte integrante della strategia digitale dell'UE e mira a promuovere l'innovazione, la competitività e lo sviluppo di un ecosistema di fiducia proponendo un quadro giuridico per un'IA affidabile nonché garantendo al contempo la protezione dei valori fondamentali dell'UE (16). La regolamentazione digitale europea è entrata in una nuova fase con l'adozione del Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024, che stabilisce norme armonizzate sull'intelligenza artificiale (RIA o AI Act in breve). Il Regolamento UE 2024/1689 del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale, meglio noto come *Artificial Intelligence Act* (AI Act), è stato, poi, pubblicato sulla Gazzetta Ufficiale dell'Unione europea il 12 luglio 2024. Le disposizioni entreranno in vigore in un periodo compreso tra i 6 e i 36 mesi dalla pubblicazione del Regolamento sulla Gazzetta Ufficiale dell'Unione europea. Il Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024, è poi, entrato in vigore l'1 agosto 2024. L'obiettivo del Regolamento è creare un quadro normativo orizzontale armonizzato per lo sviluppo, l'introduzione nel mercato dell'Unione europea e l'utilizzo di prodotti e servizi di intelligenza artificiale (AI), con particolare attenzione alla gestione dei rischi per salute, sicurezza e diritti fondamentali. Sul piano della tecnica normativa, la scelta del Regolamento è stata determinata dalla sua diretta applicabilità in tutti gli Stati Membri. Infatti, l'adozione di normative nazionali divergenti in materia avrebbe potuto comportare la frammentazione del Mercato interno e diminuire la certezza del diritto per gli operatori economici.

Come è noto, i regolamenti dell'Unione europea sono norme direttamente applicabili su tutto il territorio dell'Unione. Scegliendo questa fonte normativa, il Legislatore europeo ha inteso fissare un quadro normativo omogeneo e tendenzialmente rigido per tutti gli Stati membri. Il Regolamento in materia di intelligenza artificiale – entrato in vigore il ventesimo giorno successivo alla pubblicazione – diverrà applicabile in maniera progressiva, con tempi di efficacia delle sue disposizioni che variano tra i 6 e i 36 mesi. Sebbene limitata alla disciplina dell'utilizzo dell'intelligenza artificiale, si tratta di una regolamentazione destinata a incidere significativamente sulle scelte di *business* delle imprese private, sull'esercizio dell'attività amministrativa e, dunque, sulle strategie di *procurement* delle stazioni appaltanti.

Per quanto riguarda la strategia italiana per l'intelligenza artificiale va detto, poi, che a pochi giorni dalla pubblicazione dell'AI Act sulla Gazzetta ufficiale dell'Unione europea, è stato reso disponibile anche il testo della Strategia Italiana per l'intelligenza artificiale 2024-2026. Il documento, elaborato da un Comitato di esperti incaricato dal Governo, è finalizzato a supportare quest'ultimo nella definizione di una normativa nazionale e delle politiche sull'IA e, dopo un'analisi del contesto globale e del posizionamento italiano, definisce le azioni strategiche raggruppandole in quattro macroaree: Ricerca, Pubblica Amministrazione, Imprese e Formazione. La strategia propone, inoltre, un sistema di monitoraggio della relativa attuazione e un'analisi del contesto regolativo che traccia la cornice entro cui dovrà essere dispiegata.

(15) Per la Commissione europea la regolazione dell'intelligenza artificiale avrebbe dovuto basarsi su un'IA “etica, sicura e all'avanguardia”, posta al servizio dell'uomo e delle sfide cruciali del XXI secolo.

(16) La proposta di Regolamento sull'intelligenza artificiale [COM (2021) 206] è stata adottata dalla Commissione europea sulla base delle competenze attribuite all'Unione in materia di misure destinate ad assicurare l'instaurazione e il funzionamento del Mercato interno ai sensi dell'art. 114 del Tfu. L'approvazione di tale atto normativo segue la procedura ordinaria ai sensi degli artt. 294 ss. Tfu.

In concomitanza con l'adozione dell'*AI Act* da parte del Parlamento europeo, il Consiglio dei ministri n. 78 del 23 aprile 2024 ha deliberato un disegno di legge che disciplina l'uso dell'intelligenza artificiale nei settori demandati dal Regolamento all'autonomia normativa degli Stati Membri. L'iter di approvazione ha preso l'avvio con la presentazione del disegno di legge al Senato il 20 maggio 2024 (A.S. 1146). In tal modo anche il Governo italiano ha presentato uno schema di disegno di legge (ddl) recante disposizioni e delega al Governo in materia di intelligenza artificiale, con l'intento di individuare criteri regolatori in grado di equilibrare il rapporto tra le opportunità offerte dalle nuove tecnologie con i rischi legati ad un uso improprio. Come si legge nel Comunicato Stampa del Consiglio dei ministri le norme intervengono in cinque ambiti: la strategia nazionale, le autorità nazionali, le azioni di promozione, la tutela del diritto di autore, le sanzioni penali. Si prevede, inoltre, una delega al governo per adeguare l'ordinamento nazionale al Regolamento UE in materie come l'alfabetizzazione dei cittadini in materia di IA (sia nei percorsi scolastici che in quelli universitari) e la formazione da parte degli ordini professionali per professionisti e operatori. La delega riguarda anche il riordino in materia penale per adeguare reati e sanzioni all'uso illecito dei sistemi di IA (17).

12. La regolamentazione a livello europeo dell'intelligenza artificiale: il Regolamento (UE) 2024/1689 del Parlamento europeo

Come si è già avuto modo di accennare, il 13 marzo 2024 il Parlamento europeo ha adottato la propria posizione in prima lettura relativamente al Regolamento sull'intelligenza artificiale (IA), il c.d. *Artificial Intelligence Act* (AI Act), al fine di proteggere i diritti fondamentali, promuovendo l'innovazione, con particolare attenzione alle piccole e medie imprese, comprese le start-up e assicurando al contempo all'Europa un ruolo guida nell'adozione di un'IA affidabile ed etica.

Il Regolamento, la cui proposta era stata presentata dalla Commissione il 21 aprile 2021 come parte integrante della strategia digitale dell'UE, mira a promuovere l'innovazione, la competitività e lo sviluppo di un ecosistema di fiducia, proponendo un quadro giuridico per un'IA affidabile nonché garantendo al contempo la protezione dei valori fondamentali dell'UE. Sul piano della tecnica normativa, la scelta del Regolamento è stata determinata dalla sua diretta applicabilità in tutti gli Stati Membri. Infatti, l'adozione di normative nazionali divergenti in materia avrebbe potuto comportare la frammentazione del Mercato interno e diminuire la certezza del diritto per gli operatori economici.

Il Regolamento mira a supportare la diffusione delle tecnologie IA in diversi settori sociali ed economici, garantendo al contempo un livello uniforme di tutela degli interessi pubblici in materia di salute, sicurezza e diritti fondamentali in tutto il territorio dell'Unione europea. Sul piano del riparto delle competenze, il Regolamento non pregiudicherà l'autonomia normativa degli Stati Membri in materia di sicurezza nazionale, ricerca e in settori che non interferiscono con il Mercato interno. Non si applicherà infatti ai sistemi di IA che hanno esclusivamente scopi militari, di difesa o di sicurezza nazionale, nonché alle attività di ricerca, sviluppo e prototipazione che precedono l'immissione sul mercato o a persone che utilizzano l'IA per motivi non professionali.

L'approccio adottato dal legislatore europeo è antropocentrico e garantisce che i sistemi di IA siano sviluppati e utilizzati come strumenti al servizio delle persone, nel rispetto della dignità e dell'autonomia. L'ambito di applicazione soggettivo coinvolge l'intera catena di valore dell'IA, dai fornitori ai distributori, indipendentemente dalla collocazione della loro sede sul territorio dell'UE, purché l'output prodotto dal sistema di IA sia utilizzato nell'UE.

Il Regolamento è frutto della strategia dell'Unione Europea (UE) di forgiare un quadro legislativo europeo che preservi i valori fondamentali senza limitare la competitività europea. Questo testo normativo è il risultato di un accurato processo, che cerca di conciliare le aspirazioni della libertà imprenditoriale e la promozione del progresso tecnologico con gli imperativi del rispetto dei diritti fondamentali e della garanzia di sicurezza nei sistemi di AI.

Lo scopo del Regolamento è quello di fornire certezza giuridica attraverso un quadro unificato *ex ante* applicabile a tutti i sistemi di intelligenza artificiale immessi sul mercato dell'UE, e lo fa con diversi livelli di obblighi in base al grado di rischio che ogni sistema presenta per la salute, la sicurezza e i diritti fondamentali dei cittadini. Per lo standard, il rischio è "la combinazione della probabilità che si verifichi un danno e la gravità di tale danno" [Art. 3(2) Regolamento sull'intelligenza artificiale]. Questo approccio basato sul rischio porta il Regolamento sull'IA a proibire alcuni usi considerati di rischio inaccettabile nell'UE, a imporre obblighi rigorosi per gli usi ad alto rischio, a stabilire obblighi di trasparenza per alcuni sistemi a rischio limitato e a lasciare non regolamentati i restanti sistemi a rischio minimo. Pertanto, l'intensità normativa che incorpora è adattata al livello di rischio dello specifico sistema di AI sulla salute, la sicurezza e i diritti fondamentali dell'utente. È molto diverso dalla regolamentazione di un'auto autonoma o del correttore ortografico di Word, per esempio. Il regolamento mira a

(17) Attualmente, quindi, la regolazione dell'intelligenza artificiale si basa su due proposte di atti normativi, e precisamente, il Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 a livello europeo, e il disegno di legge sull'intelligenza artificiale di iniziativa governativa (A.S. 1146) a livello nazionale, entrambi al momento non vigenti e in attesa di approvazione da parte dei competenti organi legislativi. Cfr. su entrambi i documenti normativi il Focus del Dipartimento per il programma di Governo della Presidenza del Consiglio dei ministri, "Analisi del quadro normativo in materia di intelligenza artificiale (D.D.L. IA e Regolamento UE su IA)".

non rallentare le iniziative commerciali innovative, garantendo al contempo i diritti fondamentali delle persone. L'adozione del Regolamento sull'intelligenza artificiale è accompagnata da impatti economici positivi, che influenzano direttamente le imprese europee. Ponendo una solida base per la regolamentazione dell'intelligenza artificiale in base al livello di rischio specifico, l'UE mira a creare un ambiente di fiducia per gli utenti e i fornitori di soluzioni di AI. Questa fiducia è essenziale per stimolare gli investimenti e la ricerca in un settore così competitivo. A ciò si aggiunge l'incorporazione di ambienti di test controllati, le *sandbox*, in modo che la conformità normativa non ostacoli l'innovazione.

Da un punto di vista sociale, le implicazioni sono altrettanto significative. Con un'attenzione particolare alla supervisione normativa, il Regolamento sull'IA cerca di prevenire i rischi di discriminazione e di violazione della privacy. L'intelligenza artificiale, afferma il capitolo 2 del Regolamento sull'IA, deve servire i cittadini e facilitare "la protezione degli individui, delle imprese, della democrazia, dello Stato di diritto e dell'ambiente". Va, peraltro, osservato che la regolamentazione dell'intelligenza artificiale non è una preoccupazione esclusiva dell'UE. Altre giurisdizioni in tutto il mondo hanno già implementato o stanno valutando regolamenti con aspirazioni altrettanto globali. Tuttavia, il Regolamento sull'IA si distingue per il suo approccio completo e preventivo, che potrebbe stabilire nuovi standard globali, il cosiddetto "effetto Bruxelles" già raggiunto, per esempio, nella protezione dei dati. Guardando alla regolamentazione internazionale, possiamo notare che alcune nazioni sono impegnate in una regolamentazione settoriale sull'intelligenza artificiale o in un *laissez-faire*. D'altro canto, l'UE, con le sue specificità europee, mira ad assumere un ruolo guida in termini di diritti fondamentali. Il Regolamento sull'intelligenza artificiale mira ad allineare le pratiche industriali con i valori legali e democratici comuni a tutta l'Europa.

Il Regolamento sull'intelligenza artificiale è un regolamento molto complesso, con evidenti problemi di qualità normativa. A mio avviso, molte delle critiche sono giustificate. Tuttavia, non vanno trascurati i numerosi vantaggi che la nuova legislazione adottata intende apportare all'integrazione europea. La sicurezza per il futuro risiede in una legislazione efficace e adattabile nonostante i cambiamenti legali, sociali e tecnici nel tempo. Solo il tempo ci dirà se il Regolamento sull'intelligenza artificiale è riuscito ad anticipare queste sfide e a fornire meccanismi adeguati per affrontarle. Per lo meno, il Regolamento è un buon punto di partenza nella crescente necessità politica di regolamentare l'IA in un'ampia gamma di settori. E spetterà al CISO partecipare al processo di implementazione dei sistemi di AI in azienda, per garantire che non solo siano conformi ai requisiti di cybersecurity, ma che contribuiscano davvero a soddisfare le esigenze reali dell'organizzazione in questione.

Sul piano organizzativo il Regolamento prevede l'istituzione di un Ufficio europeo per l'IA, responsabile del coordinamento delle autorità nazionali competenti per la sorveglianza sull'attuazione delle nuove norme e l'applicazione di sanzioni amministrative. L'entità di queste è definita nel Regolamento e varia in base alla gravità e alla natura dell'infrazione nonché alla dimensione e al fatturato dell'operatore economico responsabile al fine di garantirne la capacità dissuasiva. Gli Stati Membri possono prevedere ulteriori sanzioni di natura penale per garantire la tutela degli interessi che possono essere offesi da un uso illecito dell'intelligenza artificiale.

Particolare attenzione viene posta, nel Regolamento, alla classificazione dei sistemi di intelligenza artificiale. Punto fondamentale del Regolamento (art. 3), infatti, è la definizione di sistema di IA (18) come di *"un sistema basato su una macchina progettato per funzionare con diversi livelli di autonomia e che può mostrare adattività dopo l'implementazione e che, per obiettivi espliciti o impliciti, deduce, dall'input che riceve, come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali"*. Seguendo un approccio "basato sul rischio" che l'IA può rappresentare per i diritti, la sicurezza e il buon andamento del mercato unico, la nuova disciplina stabilisce obblighi differenziati articolati su quattro livelli:

- a) rischio inaccettabile;
- b) rischio alto;
- c) rischio limitato;
- d) rischio minimo.

I sistemi di IA che determinano un rischio inaccettabile per la sicurezza e i diritti delle persone sono vietati e dovranno essere eliminati dagli Stati Membri entro sei mesi dall'entrata in vigore del Regolamento. In tale ambito rientrano i sistemi che sfruttano la vulnerabilità delle persone o tecniche deliberatamente manipolative, che consentono di attribuire un "punteggio sociale" (*social scoring*), classificando le persone in base al loro comportamento sociale o alle loro caratteristiche personali, i sistemi di riconoscimento delle emozioni utilizzati sul luogo di lavoro e negli istituti scolastici, eccetto per motivi medici o di sicurezza, i sistemi di categorizzazione biometrica delle persone fisiche sulla base di dati biometrici per dedurre dati sensibili, l'estrazione non mirata di immagini facciali da internet o telecamere a circuito chiuso per la creazione o l'espansione di banche dati e i sistemi di identificazione biometrica in tempo reale in spazi accessibili al pubblico (ossia il riconoscimento facciale mediante telecamere a circuito chiuso) da parte delle forze di polizia. Relativamente a quest'ultimo divieto sono previste eccezioni quando l'utilizzo dei sistemi sia mirato al contrasto di ipotesi di reato individuate dalla

(18) Questa definizione di sistema di intelligenza artificiale viene puntualmente ripresa dall'art. 2, lett. a) del D.D.L. sull'intelligenza artificiale.

normativa, alla ricerca di specifiche vittime, alla prevenzione di attacchi terroristici o di minacce alla vita o all'incolumità delle persone.

Il Regolamento considera, poi, ad alto rischio i sistemi di IA che possono avere un impatto “sistemico”, ossia significativo sulla sicurezza delle persone o sui diritti fondamentali, come tutelati dalla Carta dei diritti fondamentali dell'UE. Prima di immettere un sistema di IA ad alto rischio sul mercato dell'UE o di farlo entrare in servizio, i fornitori dovranno dimostrare che il loro sistema è conforme ai requisiti obbligatori per un'IA affidabile in relazione a qualità dei dati, documentazione e tracciabilità, trasparenza, sorveglianza umana, accuratezza, cybersecurity e robustezza.

Il Regolamento prevede specifici obblighi di trasparenza per i sistemi che presentino un rischio limitato di manipolazione (come attraverso l'uso di chatbot). I fornitori di sistemi di IA, che generano contenuti audio, immagini, video o di testo sintetici, dovranno garantire che i risultati siano contrassegnati e rilevabili come generati o manipolati artificialmente.

I sistemi di IA a rischio minimo (come videogiochi o filtri spam) saranno esenti da obblighi, ferma restando l'adesione volontaria a codici di condotta da parte dei fornitori di tali sistemi, che rappresentano la maggioranza di quelli attualmente utilizzati in UE.

Il Regolamento contiene diverse misure a sostegno dell'innovazione. In particolare, consente la creazione di spazi di sperimentazione normativa per l'IA (*sandbox* normativi) e di prova in condizioni reali, che forniscono un ambiente controllato per testare tecnologie innovative per un periodo di tempo limitato, promuovendo in tal modo l'innovazione da parte delle imprese, delle PMI e delle start-up. Gli Stati Membri dovranno istituire almeno un *sandbox* normativo sull'IA a livello nazionale.

13. *La regolamentazione a livello nazionale dell'intelligenza artificiale: il disegno di legge recante “Disposizioni e delega al Governo in materia di intelligenza artificiale”, approvato dal Consiglio dei ministri del 23 aprile 2024 e presentato al Senato il 20 maggio 2024 (A.S. 1146)*

Come già si è detto, in concomitanza con l'adozione dell'AI Act da parte del Parlamento europeo, il Consiglio dei ministri n. 78 del 23 aprile 2024 ha approvato un disegno di legge che disciplina l'uso dell'intelligenza artificiale nei settori demandati dal Regolamento all'autonomia normativa degli Stati Membri (19).

L'iter di approvazione ha preso l'avvio con la presentazione del disegno di legge al Senato il 20 maggio 2024 (A.S. 1146). Il testo del disegno di legge prevede le linee guida per l'utilizzo dell'IA in vari ambiti, dall'economia alla sicurezza digitale, mettendo al centro la protezione dei cittadini e dei valori democratici. Il nostro Paese, quindi, ha subito raccolto l'impegno formulato nel Regolamento europeo sull'intelligenza artificiale di armonizzare ed adeguare la legislazione nazionale a quella europea entro dodici mesi dall'approvazione dell'AI Act (20). Attraverso la disciplina prevista dal disegno di legge approvato dal Consiglio dei ministri, che si articola in 26 articoli, il nostro Paese si propone l'obiettivo di promuovere il pieno sfruttamento delle potenzialità dell'intelligenza artificiale mediante un utilizzo più responsabile, senza compromettere la sicurezza e i diritti individuali.

In particolare, il disegno di legge individua criteri regolatori capaci di riequilibrare il rapporto tra le opportunità che offrono le nuove tecnologie e i rischi legati al loro uso improprio, al loro sottoutilizzo o al loro impiego dannoso. Inoltre, introduce norme di principio e disposizioni di settore che, da un lato, promuovano l'utilizzo delle nuove tecnologie per il miglioramento delle condizioni di vita dei cittadini e della coesione sociale e, dall'altro, forniscano soluzioni per la gestione del rischio fondate su una visione antropocentrica. In quest'ottica – come si legge nello stesso comunicato stampa della Presidenza del Consiglio dei ministri a margine dell'approvazione – il disegno di legge non si sovrappone al Regolamento europeo sull'intelligenza artificiale approvato lo scorso 13 marzo dal Parlamento Europeo, di prossima emanazione, ma ne accompagna il quadro regolatorio in quegli spazi propri del diritto interno, tenuto conto che il regolamento è impostato su un'architettura di rischi connessi all'uso della intelligenza artificiale (IA) (21).

(19) Su proposta del Presidente Giorgia Meloni e del Ministro della giustizia Carlo Nordio, il Consiglio dei ministri ha approvato, in data martedì 23 aprile 2024, “*un disegno di legge per l'introduzione di disposizioni e la delega al Governo in materia di intelligenza artificiale*”. Con tale provvedimento, il governo italiano si propone di armonizzare la legislazione nazionale a quella europea entro dodici mesi dall'approvazione dell'AI Act.

(20) L'Italia è dunque in prima linea nell'adozione responsabile dell'AI. Come affermato da Alessio Butti, Sottosegretario alla Presidenza del Consiglio dei Ministri con delega all'Innovazione, “*Desideriamo riconfermare come quello italiano sia il primo governo che, subito dopo l'approvazione dell'AI Act, legifera in materia di intelligenza artificiale, peraltro con uno strumento che è il disegno di legge e non il decreto legge, ancorché molti avessero ravvisato i requisiti di necessità e urgenza per trattare un tema così delicato, proprio perché il presidente del Consiglio desidera che vi sia un confronto serrato a livello parlamentare*”. Come evidenziato dallo stesso Butti, inoltre, tale disegno di legge è un provvedimento atteso anche dagli altri Paesi del G7.

(21) Proprio sui rischi era intervenuto anche il Sottosegretario alla Presidenza del Consiglio, Alfredo Mantovano, affermando che “*L'AI concettualmente non porta nulla di nuovo rispetto alla disinformazione dei regimi totalitari, ma può aumentare a dismisura il potenziale offensivo, influenzando sul processo di manifestazione della libertà democratica. Il tema della proporzione dei controlli è quindi centrale*”.

14. *Le linee guida del disegno di legge sull'intelligenza artificiale approvato dal Consiglio dei ministri il 23 aprile 2024*

Volendo in questa sede riepilogare brevemente le linee guida del disegno di legge approvato dal Consiglio dei ministri il 23 aprile 2024, va rilevato che – come si legge ancora nel comunicato stampa del Consiglio dei Ministri – il testo del disegno di legge prevede che “il ciclo di vita dei sistemi e dei modelli di intelligenza artificiale debba basarsi sul *rispetto dei diritti fondamentali e delle libertà* dell’ordinamento italiano ed europeo oltre che sui principi di *trasparenza, proporzionalità, sicurezza, valorizzazione anche economica del dato, protezione dei dati personali, riservatezza, robustezza, accuratezza, non discriminazione, parità dei sessi e sostenibilità*”.

In particolare, il disegno di legge disciplina l’utilizzo dell’intelligenza artificiale in vari ambiti, quali, a titolo esemplificativo, la sanità e la disabilità, il lavoro, la pubblica amministrazione, l’attività giudiziaria, la cybersicurezza, la tutela degli utenti e in materia di diritto d’autore, la disciplina penale.

Con riferimento alla sanità e alla disabilità il disegno di legge stabilisce che in caso di utilizzo dell’intelligenza artificiale in ambito sanitario sarà obbligatorio informare i cittadini sull’utilizzo della tecnologia. Inoltre, sono dichiarati di interesse pubblico le ricerche e le sperimentazioni dei sistemi di AI in ambito sanitario eseguite. L’*Artificial Intelligence*, supporterà inoltre la cura e l’assistenza territoriale.

Per ciò che riguarda il lavoro, il disegno di legge prevede, poi, che l’AI può essere impiegata per migliorare le condizioni di lavoro, purché si rispetti il principio di equità e non discriminazione. È inoltre prevista l’istituzione di un Osservatorio sull’adozione dell’AI. Infine, per quanto riguarda le professioni intellettuali, “*il pensiero critico umano debba sempre risultare prevalente rispetto all’uso degli strumenti di intelligenza artificiale*”.

Con riferimento all’utilizzo dell’intelligenza artificiale nell’attività delle pubbliche amministrazioni e nella erogazione dei servizi pubblici, il disegno di legge prevede di regolare l’impiego dell’IA all’interno delle pubbliche amministrazioni garantendo l’efficienza amministrativa e rispettando i principi di autodeterminazione e responsabilità umana.

Per quanto riguarda, poi, l’attività giudiziaria, il disegno di legge prevede che l’intelligenza artificiale può essere uno strumento di supporto all’interno dell’amministrazione della giustizia. Spetta al magistrato “*la decisione sull’interpretazione della legge, la valutazione dei fatti e delle prove e sull’adozione di ogni provvedimento inclusa la sentenza*” (22).

Per ciò che riguarda la cybersicurezza nazionale, il disegno di legge prevede che l’Agenzia per la cybersicurezza nazionale promuove l’IA quale strumento per rafforzare la cybersicurezza nazionale. Al riguardo si prevede, in particolare, una disciplina specifica della strategia nazionale, delle autorità nazionali e delle azioni di promozione. Nel disegno di legge si prevede, infatti, l’introduzione della Strategia nazionale per l’IA, i cui risultati verranno riportati alle camere ogni anno. Sono inoltre introdotte le Autorità nazionali per l’intelligenza artificiale, quali l’Agenzia per l’Italia digitale (AgID) e l’Agenzia per la cybersicurezza nazionale (ACN). Queste assumeranno quindi un ruolo fondamentale nel garantire il rispetto delle normative e nell’istituire spazi di sperimentazione per promuovere l’innovazione. Sono previste anche misure a sostegno dei giovani, in particolare per l’acquisizione di nuove competenze in materia di utilizzo dell’IA negli istituti di scuola superiore. Nel complesso, nel settore dell’intelligenza artificiale, ma anche in quello della cybersicurezza e del *quantum computing* delle telecomunicazioni e delle tecnologie per queste abilitanti, è previsto l’investimento complessivo di 1 miliardo di euro. A questo si aggiungono ulteriori investimenti provenienti da fondi di CDP Venture Capital.

Nel disegno di legge non manca una specifica disciplina della tutela degli utenti e del diritto d’autore. Con il disegno di legge sull’IA si mira anche a “*favorire l’identificazione e il riconoscimento* dei sistemi di intelligenza artificiale nella *creazione di contenuti testuali, fotografici, audiovisivi e radiofonici*”. Allo stesso tempo, ci si propone di tutelare il diritto d’autore per quanto riguarda le opere generate con l’aiuto dell’intelligenza artificiale. Per ciò che riguarda, infine, la disciplina penale, è previsto un aumento della pena per i reati commessi con gli strumenti di IA, con un’ulteriore aggravante per l’alterazione dei risultati riguardanti competizioni elettorali.

(22) Sul tema specifico dell’utilizzo dell’intelligenza artificiale nell’amministrazione della giustizia è intervenuto anche il Ministro della giustizia, Carlo Nordio, affermando in conferenza stampa al termine del Consiglio dei ministri che “*ci troviamo di fronte a una vera evoluzione di cui non conosciamo gli esiti. Sappiamo che la tecnologia avanza più velocemente della legge ma noi cerchiamo di prevenirne le conseguenze. La tecnologia, come la IA non è né buona né cattiva, è neutrale. Lo dico perché si è diffuso il pensiero che l’IA possa sostituire un domani l’attività giurisdizionale. Un primo intervento è quello dell’articolo 14: Sistemi di IA sono usati esclusivamente per l’organizzazione e la semplificazione del lavoro giudiziario e la ricerca giurisprudenziale dottrinale. È sempre riservata al magistrato la decisione sull’interpretazione della legge, sulla valutazione dei fatti e delle prove e sull’adozione di ogni provvedimento. Cerchiamo di coglierne l’opportunità ma senza condizionare il magistrato*”.

15. *Il contenuto del disegno di legge sull'intelligenza artificiale approvato dal Consiglio dei ministri il 23 aprile 2024*

Come già si è avuto modo di anticipare nelle pagine precedenti, l'obiettivo del D.D.L. è la promozione di “un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a coglierne le opportunità” (art. 1 ddl) e a migliorare le condizioni di vita dei cittadini e la coesione sociale.

In particolare, il comma 2 dell'art. 1 prevede espressamente che tutte le norme devono essere interpretate e applicate in conformità al diritto dell'Unione europea. Il disegno di legge sull'intelligenza artificiale è la prima proposta legislativa in Europa tesa ad adottare le disposizioni del nuovo AI Act. Esso introduce criteri regolatori, norme di principio e di settore, che promuovono l'utilizzo delle nuove tecnologie prevedendo, al contempo, misure in grado di contenere il rischio connesso al loro uso improprio o dannoso. Le norme intervengono nei cinque ambiti della strategia nazionale, e cioè: a) autorità nazionali; b) azioni di promozione; c) tutela del diritto di autore; d) sanzioni penali; e) delega al Governo per l'adeguamento della normativa nazionale al Regolamento europeo e per la definizione della disciplina in caso di uso dell'intelligenza artificiale per finalità illecite.

Il testo si compone di 26 articoli, che disciplinano l'integrazione dell'IA in settori critici quali sanità (art. 7) e lavoro (art. 10), informazione e riservatezza dei dati personali (art. 4), sviluppo economico (art. 5), professioni intellettuali (art. 12), attività giudiziaria (art. 14), investimenti nel settore con un'autorizzazione di spesa di 1 miliardo di euro (art. 21), tutela degli utenti (art. 23), diritto d'autore, per la disciplina specifica delle opere create con l'ausilio dell'IA (art. 24) e, infine, tutela penale (art. 25), con l'introduzione di una circostanza aggravante per i reati commessi mediante l'impiego di sistemi di intelligenza artificiale, di circostanze aggravanti speciali per determinati reati e l'introduzione di una nuova fattispecie penale. Assumono particolare rilievo le norme che dettano i principi generali che devono presiedere tutto il ciclo di vita dei sistemi e dei modelli di IA, dalla fase della ricerca, sperimentazione e sviluppo fino alla fase dell'applicazione e dell'utilizzo (artt. 3, 4, 5, 6). Il disegno di legge attribuisce alla Presidenza del Consiglio dei ministri un ruolo centrale nella definizione e attuazione della strategia nazionale per l'IA (art. 17), individua nell'Agenzia per l'Italia digitale (AgID) e nell'Agenzia per la cybersicurezza nazionale (ACN) le Autorità nazionali per l'IA (art. 18), delega il Governo ad adottare i decreti legislativi necessari per l'adeguamento della normativa nazionale al Regolamento europeo del 13 marzo 2024, in particolare nel settore dell'alfabetizzazione in materia di IA nei percorsi scolastici e universitari e a definire la disciplina nei casi di uso illecito dell'IA (art. 22).

16. *Il grande limite della regolazione sull'utilizzo dell'intelligenza artificiale in ambito europeo e nazionale. Riflessioni conclusive e prospettive evolutive*

L'ingresso dell'IA nel settore pubblico rappresenta, con ogni probabilità, un processo senza ritorno. Soprattutto se tale sviluppo avverrà in modo “controllato” – all'interno di un *framework* normativo di carattere generale e condiviso all'interno dell'UE, quale si propone di essere la recente Proposta UE di regolamento dell'IA – questa tecnologia potrà apportare importanti benefici alla capacità di governo delle autorità pubbliche e, di riflesso, anche alle collettività amministrate.

Nonostante il forte ritardo, sia in termini di regolazione che di ricerca e di implementazione, rispetto ad altri Paesi europei come la Francia e la Germania, l'Italia può ancora recuperare in tempo il distacco. Il Programma Strategico intelligenza artificiale (IA) 2022-2024 ha consapevolezza di ciò e individua precise linee direttrici da perseguire già nell'immediato futuro. Affinché, però, tale documento non resti solo un “libro dei sogni”, è indispensabile un intervento del legislatore che – anche sfruttando le risorse del Pnrr – favorisca, direzionandolo, l'utilizzo di tecniche di intelligenza artificiale da parte delle pubbliche amministrazioni.

Il fatto che sia a livello europeo che a livello nazionale ci si sia dotati, o comunque è in atto un processo di regolazione dell'utilizzo dell'intelligenza artificiale va salutato con grande interesse ed è estremamente positivo per evitare che l'intelligenza artificiale venga utilizzata per finalità non positive per l'umanità.

Occorre, tuttavia, osservare che il grande limite di queste forme di regolazione non sono vincolanti a livello universale, e quindi non sono in grado che in Paesi diversi dall'Europa o dal nostro Paese l'utilizzo dell'intelligenza artificiale possa essere sottomesso e finalizzato ad interessi diversi da quello della tutela dei diritti secondo le regole previste ed approvate in ambito europeo o nazionale. Ed allora la regolazione europea o nazionale può rivelarsi insufficiente per tutelare i diritti dalla crescente espansione dell'utilizzo dell'intelligenza artificiale in ogni settore della società moderna.

In considerazione di questo grande limite della regolazione sull'utilizzo dell'intelligenza artificiale in ambito europeo e nazionale, occorre, quindi, fare appello ad altri valori e ad altre considerazioni che possano coinvolgere a livello universale e globale tutti i Grandi della Terra e impegnare a livello di regolazione e di osservanza tutti i più grandi Paesi del mondo.

Volendo, quindi, trarre qualche considerazione di sintesi e immaginare quale potrebbe esse nei prossimi anni l'incidenza dell'utilizzo dell'intelligenza artificiale sul futuro dell'umanità, appare quanto mai utile ed opportuno richiamare quanto affermato da Papa Francesco in occasione del G7 fra i Grandi della Terra svoltosi in Puglia lo scorso 14 giugno 2024. Ebbene, in quella occasione Papa Francesco, rivolgendosi ai Grandi della Terra proprio

con riferimento al tema dell'intelligenza artificiale, ha avuto modo di osservare che “(...) occorre rimettere al centro la dignità della persona in vista di una proposta etica condivisa. La stagione di innovazione tecnologica che stiamo attraversando, – ha affermato Papa Francesco – si accompagna a una particolare e inedita congiuntura sociale: sui grandi temi del vivere sociale si riesce con sempre minore facilità a trovare intese. Anche in comunità caratterizzate da una certa continuità culturale, si creano spesso accesi dibattiti e confronti che rendono difficile produrre riflessioni e soluzioni politiche condivise, volte a cercare ciò che è bene e giusto. Oltre la complessità di legittime visioni che caratterizzano la famiglia umana, emerge un fattore che sembra accomunare queste diverse istanze. Si registra come uno smarrimento o quantomeno un'eclissi del senso dell'umano e un'apparente insignificanza del concetto di dignità umana. Sembra che si stia perdendo il valore e il profondo significato di una delle categorie fondamentali dell'Occidente: la categoria di persona umana. Ed è così che in questa stagione in cui i programmi di intelligenza artificiale interrogano l'essere umano e il suo agire, proprio la debolezza dell'ethos connesso alla percezione del valore e della dignità della persona umana rischia di essere il più grande vulnus nell'implementazione e nello sviluppo di questi sistemi. Non dobbiamo dimenticare infatti – ha affermato ancora Papa Francesco – che nessuna innovazione è neutrale. La tecnologia nasce per uno scopo e, nel suo impatto con la società umana, rappresenta sempre una forma di ordine nelle relazioni sociali e una disposizione di potere, che abilita qualcuno a compiere azioni e impedisce ad altri di compierne altre. Questa costitutiva dimensione di potere della tecnologia include sempre, in una maniera più o meno esplicita, la visione del mondo di chi l'ha realizzata e sviluppata” (23). “Questo vale – ha aggiunto Papa Francesco – anche per i programmi di intelligenza artificiale. Affinché questi ultimi siano strumenti per la costruzione del bene e di un domani migliore, debbono essere sempre ordinati al bene di ogni essere umano. Devono avere un'ispirazione etica. La decisione etica, infatti, è quella che tiene conto non solo degli esiti di un'azione, ma anche dei valori in gioco e dei doveri che da questi valori derivano (...). Per questo va salutata con favore la firma a Roma, nel 2020, della Rome Call for AI Ethics (24) e il suo sostegno a quella forma di moderazione etica degli algoritmi e dei programmi di intelligenza artificiale che ho chiamato “algoretica”. In un contesto plurale e globale, in cui si mostrano anche sensibilità diverse e gerarchie plurali nelle scale dei valori, sembrerebbe difficile trovare un'unica gerarchia di valori. Ma nell'analisi etica possiamo ricorrere anche ad altri tipi di strumenti: se facciamo fatica a definire un solo insieme di valori globali, possiamo però trovare dei principi condivisi con cui affrontare e sciogliere eventuali dilemmi o conflitti del vivere”.

* * *

(23) Cfr. Intervento di Papa Francesco in occasione della Sessione sull'intelligenza artificiale del Forum Intergovernativo del G7 fra i Grandi della Terra svoltosi a Borgo Egnazia, Puglia, 14 e 15 giugno 2024.

(24) Cfr. Papa Francesco, *Discorso ai partecipanti alla Plenaria della Pontificia Accademia per la Vita*, 28 febbraio 2020.

INTELLIGENZA ARTIFICIALE E RICERCA DEL PRECEDENTE: UN AIUTO A TRANSIGERE?

di Giovanna Capilli (*)

Abstract: L'Autrice esamina l'applicazione dell'Intelligenza Artificiale (IA) nell'ambito giuridico, concentrandosi principalmente sulla possibilità da parte dei giuristi di elaborare soluzioni transattive o conciliative tramite l'IA generativa addestrata con precedenti rilevanti sottolineando l'importanza di un attento addestramento dell'IA e del ruolo del giurista nell'adattare le soluzioni proposte in un contesto privo di schemi predeterminati.

The Author analyzes the application of Artificial Intelligence (AI) in the legal domain, with a particular focus on the ability of legal experts to generate resolution proposals using generative AI trained on relevant precedents. The analysis emphasizes the importance of careful AI training and the role of the legal experts in adapting the proposed solutions to non-standardized scenarios.

Sommario: 1. *Precisazioni preliminari sull'indagine da svolgere.* – 2. *L'intelligenza artificiale: un alleato per i giuristi.* – 3. *Addestramento dell'IA e prompt engineering per una sistema efficiente.* – 4. *Conclusione: un nuovo ruolo per i giuristi.*

1. *Precisazioni preliminari sull'indagine da svolgere*

L'utilizzo dell'IA in ambito giuridico (1) costituisce una delle grandi sfide da affrontare nel prossimo futuro e ciò sia quale supporto al lavoro degli operatori del diritto e dei tribunali che potrebbero essere agevolati nell'organizzazione dei fascicoli, nella redazione dei provvedimenti, nelle ricerche giurisprudenziali, e tanto altro, sia per prevenire il contenzioso agevolando la risoluzione conciliativa della lite.

Il presente lavoro non si occuperà del tema della decisione robotica, algoritmica o predittiva, per il quale vi è già molta letteratura, ma piuttosto della possibilità di sfruttare l'IA per l'elaborazione delle decisioni esistenti, che più precisamente chiameremo “precedenti”, al fine di ottenere risultati utili ai fini transattivi o conciliativi.

Il “precedente” è qualcosa che viene prima, che è già stato, è trascorso (2), ma che comunque continua ad esercitare un'influenza significativa sia sul presente che sul futuro, cristallizza l'esistente e cerca di assicurare una certa

(*) I. Prof. avv. G. Capilli è prof.ssa ord. di Diritto privato presso l'Università San Raffaele di Roma.

II. Il *paper* è frutto dell'integrazione e rivisitazione, con note, della Relazione al Convegno “Responsabilità del medico e diritto sanitario: questioni aperte” svoltosi presso la Corte dei conti il 15 giugno 2023.

(1) Su questo tema, la letteratura è piuttosto cospicua: Barbaro, *Uso dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?*, in *Questione giust.*, fasc. 4, 2018, scaricabile al link: <https://www.questionegiustizia.it/rivista/articolo/uso-dell-intelligenza-artificiale-nei-sistemi-giud-591.php>; Battelli, *Giustizia predittiva, decisione robotica e ruolo del giudice*, in *Giust. Civ.*, 2020, 2, p. 281 ss; Luciani, *La decisione giudiziaria robotica*, in *Riv. assoc. it. cost.*, 2018, 3, 872 ss; Ainora, *Smart courts: verso una nuova dimensione della giustizia*, in *Ius* 30 luglio 2021; Aa.Vv., *Decisione robotica*, a cura di Carleo A., Bologna, 2019; Giordano, *La giustizia predittiva*, in *Ius* 20 febbraio 2020; Viola, *Interpretazione della legge con modelli matematici*, Milano, 2017; Taruffo, *Judicial Decisions and Artificial Intelligence*, in *Judicial Applications of Artificial Intelligence*, a cura di G. Sartor - L.K. Branting, 1998, 207 ss; A. de la Oliva Santos, “*Giustizia predittiva*”, *interpretazione matematica delle norme, sentenze robotiche e la vecchia storia del “Justizclavier”*, RTDPC, 2019, 3, 883; Nieva-Fenoll, *Intelligenza artificiale e processo*, 2018, trad. it. P. Comoglio, Giappichelli, 2019; Gabellini, *Algoritmi decisionali e processo civile: limiti e prospettive*, RTDPC, 2022, 1, 59 ss; Libertini, Maureri, Vincenti, *Intelligenza artificiale e giurisdizione ordinaria. Una ricognizione delle esperienze in corso*, in *Intelligenza artificiale e diritto: una rivoluzione?*, a cura di A. Pajno, F. Donati, A. Perrucci, vol. II, Amministrazione, responsabilità, giurisdizione, il Mulino, 2022, 496 ss; Larret-Chahine, *La justice Prédictive*, in *La decisione nel prisma dell'intelligenza artificiale*, a cura di E. Calzolaio, Wolters Kluwer, 2020, 161 ss; De Pasquale, *La giustizia predittiva in Francia: il trattamento DataJust*, *Judicium*, 2021; Martinez Gutierrez, *Intelligenza artificiale, algoritmi e automazione nella giustizia. Proposte per un'effettiva implementazione*, *Processo*, 2021, 3, 442 ss; Gabellini, *La «comodità nel giudicare»: la decisione robotica*, in *Riv. trim. dir. e proc. civ.*, 2019, 1305; Morelli, *Intelligenza artificiale, Essere avvocati nell'era di ChatGPT*, Maggioli 2024.

(2) In dottrina sul precedente si veda: Rordorf, *Nomofilachia e precedente giudiziario. Il precedente nella giurisprudenza*, in *Foro It. Vol.* 142, 9, 2017, pp. 277-285; Maniaci, *Il precedente giudiziale in Italia: verso lo stare decisis?*, in *Riv. critica dir. priv.*, 2014, pp. 567-590; Serio, *La rilevanza del fatto nella struttura del precedente giudiziario inglese*, in *Eur. dir. priv.*, 2013, 357; Mazzamuto, *La Suprema Corte ed il vincolo del precedente*, in *Eur. dir. priv.*, fasc. 4, 2020, p. 1089 ss.

Al ruolo del giudice e alla funzione interpretativa è dedicato anche il fascicolo “*Il giudice e la legge*”, in *Questione giust.*, 2016, fasc. 4 scaricabile dal link: <https://www.questionegiustizia.it/rivista/2016-4.php>; nell'editoriale Rordorf evidenzia afferma: “Il compito del giudice (al pari di qualsiasi altro interprete della legge) consiste perciò oggi soprattutto nel testare i limiti di elasticità della regola dettata dal legislatore, nel comprendere quale sia, di volta in volta, lo spazio interpretativo che nella situazione data l'ordinamento nelle sue varie articolazioni gli offre, e di saper motivare le scelte che egli compie nello spazio segnato da quei limiti, in modo da favorire il formarsi di indirizzi giurisprudenziali coerenti (nel che consiste, come

prevedibilità delle decisioni future (3); nel contesto giuridico ci si riferisce al “precedente giudiziario” per indicare un caso già deciso simile a quello sul quale occorre ancora decidere (4). Questa analogia sostiene l’idea che situazioni simili dovrebbero essere decise allo stesso modo e ciò anche per garantire coerenza ed equità nel sistema giuridico.

Per valutare se un precedente può essere utile nella risoluzione di un altro caso, occorre confrontare le situazioni coinvolte per accertare se la stessa regola giuridica sia stata applicata a casi simili e, perciò, sia rilevante per la risoluzione del nuovo caso.

Tuttavia, è cruciale distinguere tra confronto casistico singolo e formazione di orientamenti giurisprudenziali consolidati.

Nel primo caso, infatti, siamo di fronte a situazioni singolarmente considerate per le quali occorre garantire la coerenza tra decisioni ed evitare che la vicenda successiva sia decisa in modo difforme da quella anteriori; nel secondo caso si parla di orientamenti giurisprudenziali (più o meno) consolidati – che possono derivare da una serie di precedenti o essere anche l’esito di un caso singolo particolarmente significativo (*leading case*) – e che contribuiscono a dare coerenza (per quanto possibile) all’ordinamento nel suo complesso.

Indipendentemente dall’approccio seguito nei vari ordinamenti giuridici, sia di *common law* che di *civil law*, il valore del precedente solitamente non è vincolante, bensì soltanto persuasivo, cosicché non vi è dubbio possa fornire un quadro utile per raggiungere una transazione o una conciliazione o quantomeno indirizzarla (5).

Il fenomeno dell’inflazione dei provvedimenti giurisdizionali, soprattutto di legittimità, tuttavia, rappresenta una criticità poiché produce un’enorme quantità di precedenti e rende complesso il compito di identificare linee di sviluppo giurisprudenziali coerenti e chiare. Aspetto questo che, nell’ambito dell’analisi dei dati da parte di sistemi di intelligenza artificiale, diventa estremamente complesso, anche in ragione del fatto che non tutti i precedenti sono ugualmente rilevanti ai fini transattivi, e si deve essere in grado di distinguere le pronunce che effettivamente sono utili e significative per il processo di transazione o conciliazione. Capacità di discernimento questa che è essenziale per garantire un utilizzo efficace del precedente nell’ambito delle negoziazioni legali.

2. L’intelligenza artificiale: un alleato per i giuristi

Chiarito l’oggetto di indagine, è importante ora esaminare come l’intelligenza artificiale si possa applicare in ambito giuridico, poiché, al netto dei rischi che questi sistemi indubbiamente presentano, le opportunità offerte dall’intelligenza artificiale sono molteplici e meritevoli di attenzione.

È cruciale cogliere queste nuove opportunità per non restare indietro rispetto all’evoluzione della tecnologia ed è indispensabile acquisire conoscenze sul funzionamento dell’intelligenza artificiale per imparare ad utilizzarla al meglio delle sue potenzialità.

già altre volte mi è capitato di dire, la funzione della moderna nomofilachia: non conservazione statica di orientamenti tradizionalmente recepiti, bensì capacità di governare l’evoluzione giurisprudenziale in funzione dei mutamenti sempre più accelerati della società e delle esigenze del sistema giuridico che essa esprime”. Franzoni, *Diritto, processo e precedente giudiziario*, in *Pol. dir.*, 2013, 415; Luchetti, *Il valore del precedente giurisprudenziale sul confine mobile tra potere legislativo e potere giudiziario*, in *Arch. civ.*, 2014, 507; Martinuzzi, *Le conseguenze giuridiche dell’inosservanza del precedente giudiziario in ambito processuale*, in *Riv. trim. dir. e proc. civ.*, 2016, 1345 Canzio, *Nomofilachia, valore del precedente e struttura della motivazione*, in *Foro it.*, 2012, V, 305.

(3) Giordano, *Il precedente nell’era digitale*, in AA, *Il diritto nell’era digitale*, Giuffrè, 2022, p. 1021 ss che, in verità, propone di usare l’intelligenza artificiale “per creare banche dati “smart” le quali, in base alle istruzioni date dall’algoritmo, siano in grado di individuare, scandagliando ogni parte delle pronunce contenute nel sistema, ciò che di esse interessa realmente per risolvere la specifica controversia”.

(4) Sulla differenza tra “precedente” e “massima” si rinvia a: Damosso, *La massima come fonte nel sistema del precedente*, in Cass. Pen., fasc. 4, 2020, p. 1708 ss, il quale nell’evidenziare che la prassi del precedente giurisprudenziale in Italia si configura come un sistema articolato e influente, che contribuisce in modo significativo alla coerenza e alla prevedibilità delle decisioni giudiziarie, conclude affermando che la massimazione riveste un ruolo cruciale nel sistema dei precedenti, poiché fornisce un importante strumento per comprendere la vasta giurisprudenza di legittimità e contribuisce sia alla formazione del precedente, attraverso la creazione di un materiale giurisprudenziale essenziale con cui gli operatori del diritto possono confrontarsi, sia alla sua interpretazione e applicazione. In sostanza, la pratica della massimazione “genera” il precedente.

(5) Non potendo approfondire la distinzione tra precedente e principio di diritto, in questo contesto ci si limiterà sinteticamente ad evidenziare che tale differenza risulta più evidente nei sistemi di *civil law* in cui il principio di diritto implica la formulazione di criteri interpretativi astratti e generalizzanti che possono essere applicati a una serie di casi simili, mentre il precedente giudiziario contiene le specificità di un singolo caso.

Discorso a parte, poi, per le clausole generali, le quali richiedono al giudice di integrare la regola legale attraverso l’uso di criteri e valori extralegali. Questi criteri e valori possono essere tratti direttamente dall’esperienza comune o da altre aree di conoscenza.

Si possono descrivere due diversi approcci nell’interpretazione e nell’applicazione del diritto basati rispettivamente sulla concretezza (il precedente giudiziario che si basa sul singolo caso) e sull’astrattezza (i principi di diritto e l’ampia discrezionalità del giudice nell’utilizzo delle clausole generali), tenuto conto che ogni caso giudiziario è unico e quasi mai identico ad un altro (fatta eccezione per i casi seriali), per cui è essenziale distinguere tra i profili di similitudine che consentono il riferimento al precedente e quelli di differenziazione, che, invece, possono suggerire un approccio contrario.

Dal punto di vista definitorio l'AI Act, Regolamento (UE) 2024/1689 del 13 giugno 2024 (6), fornisce la seguente definizione:

“sistema di IA”: un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali (7).

Al momento non esistono strumenti in grado di gestire e risolvere automaticamente i casi giudiziari senza l'intervento umano, si deve quindi pensare all'IA come assistente dei professionisti del diritto e precipuamente per la prevenzione del contenzioso o per la definizione dello stesso in via alternativa al giudizio. In questo contesto, il ruolo del giurista diventa ancora più determinante: egli agisce come un facilitatore o un supervisore dell'*output* fornito dall'intelligenza artificiale, al fine di evitare situazioni di disuguaglianza o soluzioni ingiuste che potrebbero derivare dall'uso di algoritmi o sistemi di *machine learning*.

Questi strumenti sono estremamente utili in termini di velocità nell'esecuzione di diverse operazioni, come la ricerca giurisprudenziale, il suggerimento di strategie legali o la preparazione di documenti. Tuttavia, non appaiono ancora così sofisticati (8), poiché è imprescindibile il costante intervento umano per garantire la trasparenza e la correttezza nel processo decisionale.

Ciò significa che è fondamentale per qualsiasi soggetto interessato comprendere su quali dati si basa l'intelligenza artificiale per generare *output*, e quindi perché ne vengono selezionati alcuni rispetto ad altri (9).

Conoscere, quindi, quali sono i dati di addestramento, i dati di convalida o i dati di input diventa determinante per comprendere anche come si sia addivenuti ad un determinato *output* (esempio una proposta conciliativa o transattiva); per un algoritmo non è ancora possibile simulare il ragionamento giuridico, ma potrebbe essere di supporto per effettuare previsioni, identificare i rischi ed opportunità e generare raccomandazioni sulle possibili soluzioni (10), purché il momento finale sia gestito sempre e comunque dall'intervento umano, al momento, unico capace di comprendere tutte le sfumature, le peculiarità e le specificità irripetibili di ogni singolo caso.

(6) Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale), in Gazzetta Ufficiale dell'Unione Europea IT Serie L del 12 luglio 2024.

Per un commento sul nuovo Regolamento Europeo: Aa.Vv., Commentario al nuovo Regolamento europeo sull'Intelligenza Artificiale, a cura di Cassano-Tripodi, Maggioli 2024.

(7) In Italia, si segnala la proposta di legge sull'IA (disegno di Legge, atto senato n. 917 presentato il 19 ottobre 2023) in cui all'art. 1., c. 1, contenente la definizione si legge: *“Ai fini della presente Legge, si intendono per «contenuti generati da intelligenza artificiale (IA)» tutti i contenuti editoriali, compresi testi, video, immagini e voci, che sono creati, generati o sintetizzati, in tutto o in parte, da sistemi basati su intelligenza artificiale (IA), ivi compresi algoritmi di apprendimento automatizzato, cosiddetto machine learning, e reti neurali artificiali”*. Viene, quindi, definito il concetto di “contenuto generato da artificiale IA”, specificandone le componenti essenziali; si segnala altresì la proposta di legge A.C. 1514, recante *“Disposizioni per assicurare la trasparenza nella pubblicazione e diffusione di contenuti prodotti mediante sistemi di intelligenza artificiale”*. Il testo, di cui non è stato ancora avviato l'esame, consta di 3 articoli.

(8) Vedi per esempio il sistema *GiuriMatrix* (<https://www.giurimatrix.it>) che ancora presenta diverse criticità nel suo utilizzo che gli ideatori stanno cercando di superare.

(9) Nell'art. 3 dell'AI Act sono contenute le seguenti definizioni:

29) *“dati di addestramento”: i dati utilizzati per addestrare un sistema di IA adattandone i parametri che può apprendere;*

30) *“dati di convalida”: i dati utilizzati per fornire una valutazione del sistema di IA addestrato e per metterne a punto, tra l'altro, i parametri che non può apprendere e il processo di apprendimento, al fine tra l'altro di evitare lo scarso (underfitting) o l'eccessivo (overfitting) adattamento ai dati di addestramento;*

31) *“set di dati di convalida”: set di dati distinto o costituito da una partizione fissa o variabile del set di dati di addestramento;*

32) *“dati di prova”: i dati utilizzati per fornire una valutazione indipendente del sistema di IA al fine di confermarne le prestazioni attese prima della sua immissione sul mercato o messa in servizio;*

33) *“dati di input”: i dati forniti a un sistema di IA o direttamente acquisiti dallo stesso, in base ai quali il sistema produce un output;*

(10) In quest'ottica, per esempio, si pone un nuovo sistema recentemente pubblicizzato denominato *GiuriMatrix* e in cui proprio gli ideatori evidenziano nella descrizione che si tratta di un nuovo assistente legale basato sull'intelligenza artificiale che si pone quale supporto ai professionisti del diritto al fine di svolgere il loro lavoro in modo più efficiente, al fine di migliorare la qualità del loro lavoro, risparmiare tempo prezioso e aumentare la loro produttività. Un aspetto fondamentale che hanno ritenuto di inserire gli ideatori è quello di sottolineare che le risposte sono individuate nella legge con citazione della fonte e la dottrina e la giurisprudenza vengono utilizzati come mezzi per trovare la risposta.

Gabellini, *op. cit.*, in cui sono richiamati diversi sistemi quali *Ross*, *Prédicteice*, *Case Law Analytics*, *Case Cruncher Alfa*.

Pensiamo, poi, a chatGPT modello di intelligenza artificiale avanzato e sofisticato di tipo generativo, che è capace di dare soluzioni specifiche interagendo con gli utenti anche in diversi contesti e lingue grazie anche ad un generatore *hardware* per l'apprendimento profondo sviluppato da Google che utilizzano reti neurali oltre che algoritmi di apprendimento automatico.

Il Regolamento europeo sull'intelligenza artificiale, infatti, include “i sistemi di IA destinati ad essere usati da un'autorità giudiziaria o per suo conto per assistere un'autorità giudiziaria nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti, o a essere utilizzati in modo analogo nella risoluzione alternativa delle controversie” nell'ambito dell'elenco dei sistemi di IA ad alto rischio di cui all'art. 6 paragrafo 2 e nel considerando n. 61 in cui si specifica ulteriormente che “l'utilizzo di strumenti di IA può fornire sostegno al potere decisionale dei giudici o all'indipendenza del potere giudiziario, ma non dovrebbe sostituirlo: il processo decisionale finale deve rimanere un'attività a guida umana”.

3. Addestramento dell'IA e prompt engineering per un sistema efficiente

Uno degli aspetti fondamentali nell'utilizzo dei sistemi di intelligenza artificiale è rappresentato dall'addestramento (11) che comprende evidentemente la selezione delle fonti da utilizzare a tal fine ed anche un'ottima capacità di *prompt engineering* (12). In ambito giuridico, un prompt ben formulato può fare la differenza tra un'analisi accurata e una fuorviante e, quindi, può fornire *output* utili ed utilizzabili per aiutare il giurista nel proprio compito (nel caso di specie riuscire a trovare una soluzione bonaria della controversia che sia accettabile dalle parti coinvolte).

Negli Orientamenti Etici per un'IA affidabile (Ethics guidelines for trustworthy AI) (13) viene sottolineato come affidabilità e integrità dei dati siano aspetti cruciali per il funzionamento ottimale dei sistemi di intelligenza artificiale; ciò significa che i set di dati utilizzati debbano essere di alta qualità e senza distorsioni, imprecisioni o errori socialmente costruiti. Da qui la necessità di affrontare queste questioni prima di addestrare un sistema di IA con un particolare set di dati per evitare che l'introduzione di dati malevoli possa influenzare il comportamento del sistema, specialmente nei sistemi di autoapprendimento nonché la necessità di testare e documentare i processi e i set di dati utilizzati in tutte le fasi, inclusa la pianificazione, l'addestramento, la prova e la distribuzione; ciò deve valere anche per i sistemi di IA acquisiti esternamente, non sviluppati internamente.

Anche l'AI Act laddove esamina l'addestramento evidenzia che se il sistema non è addestrato con dati di alta qualità e non soddisfa i requisiti adeguati con riferimento alle sue prestazioni, alla sua accuratezza o robustezza, potrebbe comportare problemi in ordine a discriminazioni qualora non siano sufficientemente trasparenti, spiegabili e documentabili (14).

Ecco perché si prevede che i sistemi di IA ad alto rischio dovrebbero essere progettati e sviluppati in modo tale che le persone fisiche possono supervisionare il loro funzionamento, garantire che siano utilizzati come previsto e che i loro impatti vengono affrontati durante il ciclo di vita del sistema. Un'adeguata supervisione umana, quindi, è sempre necessaria.

Tornando all'importanza dell'addestramento nell'ambito dell'intelligenza artificiale è essenziale considerare la stretta correlazione tra addestramento e dominio, che rappresenta una specifica area di conoscenza o un settore di applicazione.

(11) Nel considerando n. 67 dell'AI Act si legge: “Per disporre di set di dati di addestramento, convalida e prova di elevata qualità è necessario attuare adeguate pratiche di governance e gestione dei dati. I set di dati di addestramento, convalida e prova, incluse le etichette, dovrebbero essere pertinenti, sufficientemente rappresentativi e, nella misura del possibile, esenti da errori e completi nell'ottica della finalità prevista del sistema”.

Per un approfondimento sul tema si veda anche D'Amora – Uletto, *Le modalità di realizzazione dei sistemi e la governance sui dati*, in Aa.Vv. (a cura di Cassano – Tripodi), cit., p. 507 in cui gli autori evidenziano che “al pari dell'uomo, anche il sistema di IA ha costante necessità di mantenere la propria “mente” attiva, continuando a raccogliere dati e informazioni, per migliorare la propria efficienza e per raggiungere in maniera sempre più soddisfacente l'obiettivo per cui quella tecnologia di IA è stata creata. Tuttavia, nel caso di utilizzo dell'intelligenza artificiale, è bene ricordare che la raccolta dei dati viene effettuata non da un umano bensì da un algoritmo, entità impossibile da immaginare nel contesto spazio-temporale e difficilmente definibile anche solo a livello puramente terminologico. Questa entità misteriosa raccoglierà i dati immessi nel sistema di IA e di cui il sistema di IA necessita per funzionare correttamente e li processerà ed analizzerà, al fine di svolgere l'attività per cui la tecnologia stessa è stata ideata e posta sul mercato nonché raggiungere al meglio l'obiettivo per cui è stata sviluppata”.

(12) La *prompt engineering* è la disciplina di scrivere istruzioni precise (*prompt*) per guidare l'IA nel completamento di un compito. Sul tema si veda per esempio il testo: Bonaventura Di Bello, *A.I. Pro, Progettare, creare e comunicare con l'intelligenza artificiale generativa*, Milano 2024, dove a p. 31 viene spiegato anche come imparare ad usare l'IA generativa attraverso il prompt. Analizza i diversi sistemi di IA generativa Unfer, *ChatGPT e sistemi simili*, in Aa.Vv. (a cura di Cassano-Tripodi), cit., p. 667 e ss; Morelli, cit., p. 206 che scrive “Il *prompt engineering* permette in sostanza di a) interloquire con una applicazione di *large language modelling*, indirizzando la macchina al fine di b) ottenere l'*output* più vicino a quello desiderato.... tramite il *prompt*, uomo e macchina conversano”.

(13) Il testo può essere scaricato dal seguente link: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>; nel testo è anche inserita una versione pilota di “Lista di controllo per la valutazione dell'affidabilità dell'IA”.

(14) Da ciò la distinzione anche tra sistemi IA ad alto e basso rischio, laddove i sistemi di IA ad alto rischio dovrebbero essere progettati e sviluppati in modo tale che le persone fisiche possono supervisionare il loro funzionamento, garantire che siano utilizzati come previsto e che i loro impatti vengono affrontati durante il ciclo di vita del sistema. A tal fine è necessaria un'adeguata supervisione umana.

Il contesto medico o il contesto giuridico, per esempio, sono settori specializzati con terminologia propria e problemi peculiari (15).

Nel momento in cui si deve addestrare un dominio specifico ovviamente si dovrà effettuare un processo di personalizzazione del modello per comprendere e gestire meglio le informazioni e le sfide associate a quel particolare settore.

Ciò implica la scelta delle “fonti” (o dati) con cui addestrare il sistema (16).

In base a quanto si è detto sopra, quindi, se il “dominio” è la responsabilità medica si può immaginare un addestramento focalizzato da un lato sulle questioni tipiche giuridiche che caratterizzano la responsabilità in ambito sanitario, e dall’altro anche sulle questioni tipicamente mediche al fine di poter valutare e raggruppare casi di un certo tipo che hanno ricevuto una determinata soluzione con l’obiettivo di generare come *output* soluzioni da adattare al caso concreto.

Ciò su cui occorre evidentemente porre molta attenzione è l’adozione di misure precauzionali per la progettazione e l’uso, per evitare il cosiddetto effetto “performativo” o di autorealizzazione, cioè il rischio che un sistema produca progressivamente gli stessi risultati in uscita influenzando a sua volta i produttori di informazioni in entrata.

Unitamente all’addestramento, diventa fondamentale per il giurista creare *prompt* efficaci per l’intelligenza artificiale (17), ciò significa che il *prompt* dovrà essere chiaro, conciso e specifico, perché più è dettagliato più preciso sarà il risultato dell’IA; dovrà essere formulato in modo accurato e senza ambiguità; dovrà essere pertinente al compito che si desidera che l’IA svolga (ciò significa evitare l’inserimento di informazioni non necessarie che possono confondere l’IA e generare risultati irrilevanti); dovrà essere strutturato in modo logico e sequenziale e utilizzare una terminologia legale corretta.

In settori complessi come quello della responsabilità medica, l’addestramento e la formulazione di *prompt* efficaci assumono un ruolo ancora più cruciale in ragione della peculiarità del linguaggio e della casistica esistente. Va da sé, infatti, che termini tecnici possono risultare ambigui o fuorvianti per un’IA non addestrata in modo corretto che potrebbe fraintendere il significato di parole chiave giungendo a conclusioni errate e generare output non accurati (si parla anche delle c.d. “allucinazioni”, poiché i sistemi di IA possono fornire risposte incoerenti o non veritiere magari basate su precedenti inesistenti creati dallo stesso sistema di intelligenza artificiale). (....)

Si deve evitare, poi, che l’output generato abbia un valore prescrittivo e non indicativo, in quanto in tal modo si finirebbe per seguire ciecamente la soluzione proposta senza valutarla attentamente alla luce delle specifiche circostanze del caso in esame.

Diventa, quindi, essenziale mantenere un approccio flessibile e critico nell’uso dell’intelligenza artificiale nel contesto legale, perché se a prima vista sembrerebbe allettante predisporre dei meccanismi di trattamento più o meno automatizzati della giurisprudenza per limitare contenziosi civili ripetitivi e di basso valore (Gran Bretagna, Paesi Bassi

(15) D’Elia, *Gli strumenti di intelligenza artificiale generativa nel contesto sanitario: problemi di ottimizzazione delle risorse e questioni di spiegabilità*, in Riv. it. medicina legale, 2023, 2, p. 339 ss.

(16) Per esempio, sul sito di *GiuriMatrix* si legge: “Le risposte sono individuate nella legge con citazione della fonte. La dottrina e la giurisprudenza vengono utilizzati come mezzi per trovare la risposta”: il sistema *GiuriMatrix* è un sistema deduttivo, con ottime potenzialità, sebbene ancora sia in fase iniziale di sperimentazione; tuttavia, già consente di ottenere risultati interessanti; al seguente *link* è possibile visionare un video in cui propone una transazione per la risoluzione di un caso che viene proposto: <https://www.youtube.com/watch?v=D7P6x2LsQlo>; altro dato interessante è che *GiuriMatrix* non appare discriminatorio proprio perché la sua base sono le norme e non la giurisprudenza. Un altro esempio di utilizzo di IA per la risoluzione dei conflitti è il progetto Crea <https://www.crea-project.eu> (Conflict Resolution with Equitative Algorithms), sul sito è possibile scaricare anche il manuale operativo <https://www.crea-project.eu/wp-content/uploads/2020/01/CREA-Handbook.pdf> in cui si legge che il progetto mira a introdurre nuovi meccanismi di risoluzione delle controversie come strumento di supporto per legali, mediatori e giudici, con l’obiettivo di raggiungere un accordo tra le parti e in alcuni casi, potrebbe essere utilizzato direttamente dai cittadini.

(17) Juan David Velásquez-Henao, Carlos Jaime Franco-Cardona, Lorena Cadavid-Higuaita, *Prompt Engineering: a methodology for optimizing interactions with A.I.-Language Models in the field of engineering*, <https://doi.org/10.15446/dyna.v90n230.111700>; Mohammad Aljanabi, Mohanad Ghazi Yaseen, Ahmed Hussein Ali, Mostafa Abdulghafoor Mohammed, *Prompt Engineering: Guiding the Way to Effective Large Language Models*, <https://doi.org/10.52866/ijcsm.2023.04.04.012>; Korzyński, Paweł; Mazurek, Grzegorz; Krzyżkowska, Pamela; Kurasinski, Artur, *Artificial intelligence prompt engineering as a new digital competence: Analysis of generative A.I. technologies such as ChatGPT*, Entrepreneurial Business and Economics Review; Krakow Vol. 11, fasc. 3, 2023: 25-37. DOI:10.15678/EBER.2023.110302, in cui viene messo in evidenza l’abilità di generare *prompt* come “a skill can lead to the generation of more accurate and contextually relevant results, which in turn improves the overall performance and efficiency of the AI models”.

(...) Morelli, cit., p. 192 riporta la definizione del Cambridge Dictionary sul tema dell’allucinazione: “La nuova definizione aggiuntiva di allucinazioni – spiega il Cambridge Dictionary – è: quando un’intelligenza artificiale (= un sistema informatico che possiede alcune delle qualità del cervello umano, come la capacità di produrre un linguaggio in un modo che sembra umano) ha allucinazioni, produce false informazioni”. A p. 174 viene riportato il caso di un avvocato americano sanzionato dal Tribunale di New York per aver richiamato negli atti processuali alcune sentenze inesistenti, frutto delle allucinazioni del sistema ChatGPT.

e Lettonia sono esempi di Paesi che hanno già attuato, o stanno per attuare, questo tipo di soluzione), dall'altro non si può abdicare al rispetto di trasparenza, neutralità e lealtà (18).

Se, per esempio, il sistema di IA fornisce una proposta di tipo conciliativo o transattivo, si dovrebbe prevedere la possibilità di discuterla in contraddittorio con l'aiuto di una terza parte qualificata e certificata, facendo salva in ogni caso la possibilità di accedere al giudice.

Immaginiamo allora quanto potrebbe essere utile un sistema di intelligenza artificiale addestrato su precedenti giudiziari, in cui si possono inserire gli *input* necessari per formulare una proposta basandosi su ciò che si ritiene fondamentale per la risoluzione del caso (esempio atti o documenti) cosicché il ruolo del giurista non viene ridotto, ma piuttosto amplificato e trasformato diventando il motore che avvia il processo di generazione dell'*output* in base a tutti gli elementi forniti dalle parti ma selezionati in base alla valutazione (19) di rilevanza degli stessi in relazione alla situazione da risolvere.

La necessità di una legislazione appropriata volta ad implementare l'intelligenza artificiale quale strumento di ausilio agli operatori della giustizia diventa imprescindibile, oggi ancora non ci sono strumenti né regole specifiche, anche se l'AI Act costituisce indubbiamente la base fondamentale su cui costruire regole specifiche in questo settore.

4. Conclusione: un nuovo ruolo per i giuristi

La breve analisi svolta dimostra come l'avvento dell'intelligenza artificiale nel panorama giuridico segni una trasformazione epocale, che comporterà la necessità di ridefinire radicalmente le modalità di approccio e gestione del diritto.

Sebbene una sostituzione degli uomini con le macchine appaia un'eventualità piuttosto remota, diventa necessario creare una collaborazione sinergica tra uomo e IA al fine di potenziare l'efficienza, l'accuratezza e l'accessibilità della giustizia.

Inoltre, l'IA potrà potenziare le capacità del giurista il cui ruolo, probabilmente, si andrà a modificare con il tempo, perché non sarà più solo un esperto di leggi, ma dovrà avere competenze per l'addestramento e l'interrogazione dell'IA. (20).

In sostanza, si amplificherà ciò che già il giurista effettua nel momento in cui svolge una ricerca giurisprudenziale o dottrinale con le attuali banche dati, con la differenza che dovrà sapere selezionare meglio i dati e formulare domande precise e interpretare gli output vagliandone l'affidabilità e la coerenza con il contesto giuridico.

L'intelligenza artificiale, quindi, deve essere vista come un prezioso e strategico alleato che permetterà di potenziare le nostre abilità anche nell'ottica di una migliore gestione dei contenziosi, poiché la capacità di porre le domande giuste e comprendere i limiti e le potenzialità dell'IA costituirà un'abilità distintiva per i giuristi di domani e sarà anche la chiave del loro successo.

In questo scenario, la formazione avrà un ruolo di primaria importanza, perché sarà necessario strutturare nuovi percorsi didattici in cui le competenze giuridiche tradizionali siano integrate con le conoscenze di base sull'IA, sul *machine learning* e sul *data analysis*.

La rivoluzione tecnologica si è avviata, ora tocca a noi essere intraprendenti e non lasciarci sfuggire le occasioni che offre.

* * *

(18) Per esempio: sarà calcolata la compensazione? Con quale metodo? L'algoritmo è in grado di elaborare correttamente le informazioni? La proposta può essere discussa in modo contraddittorio con l'aiuto di una terza parte qualificata e certificata? L'accesso al giudice è sempre possibile? È essenziale che principi di trasparenza, di neutralità e lealtà verso i cittadini siano integrati nei meccanismi in questione.

(19) Licci, *L'impatto dell'intelligenza artificiale sui poteri istruttori del giudice e sulla definizione dei fatti notori*, in AA, *Il diritto nell'era digitale*, Giuffrè, 2022, p. 855 ss, in cui si legge: "Con la tecnologia è possibile ricercare automaticamente i documenti potenzialmente rilevanti ai fini della decisione, replicando i risultati dell'attività umana corrispondente, secondo un processo che tuttavia non ha nulla a che vedere con la valutazione e il pensiero dell'uomo. Un sistema così congegnato è in grado, perciò, potenzialmente di garantire una maggiore efficienza e precisione rispetto all'agire umano, con un risparmio di costi e di tempi"; interessante è la valutazione del c.d. fatto notorio. Quest'ultimo, infatti, nell'analisi di cui al presente scritto potrebbe essere utilizzato ai fini dell'addestramento dell'IA.

(20) Si interroga sulle nuove frontiere delle professioni legali Morelli, cit., p. 205 in cui scrive: "Siamo alla vigilia di una nuova trasformazione nel settore professionale forense? Il dibattito è aperto, tra chi pensa che stia emergendo una nuova figura di avvocato, che dovrà integrare questa nuova competenza conversazionale, nella relazione uomo-macchina; e chi pensa che stia nascendo una nuova figura professionale da impegnare negli studi legali: il GPT Legal Prompt Engineering, appunto". A p. 210, l'A. ha inserito anche la guida ai prompt di Open AI.

Guido Alpa, nella prefazione al testo, scrive che l'intelligenza artificiale "sarà uno strumento straordinario che aiuterà i giuristi a raccogliere dati, a valutarli, a redigere gli atti, perfino a giudicare. È una prospettiva che aiuta a riflettere, e che si pone come imperativa, perché non si può arrestare lo sviluppo e l'applicazione dell'intelligenza artificiale, né invertire la rotta".

CONCLUSIONI

di Alberto Avoli (*)

Aristotele secoli fa si era proposto l'obiettivo di risolvere l'annosa questione se l'uovo fosse nato prima della gallina o viceversa e, come risaputo, aveva argomentato che l'uovo è una gallina in potenza, mentre la gallina è un uovo già sviluppato, con la conseguenza tutta concettuale che la gallina aveva preceduto l'uovo in quanto essere superiore.

Ai nostri giorni il tema potrebbe risultare ozioso e di "ovvia" veloce soluzione: prima dell'uovo e prima della gallina sarebbe certamente venuto l'algoritmo.

L'algoritmo, l'intelligenza artificiale e una serie infinita di lemmi e di acronimi, incomprensibili ai più, pervadono la nostra quotidianità, con il rischio di ingenerare l'illusione di meccanismi arcani in grado di per sé di risolvere tutti i problemi dell'umanità, portando ordine nel disordine, creando la magia dell'arte con un click, riempiendo il presente e il futuro con dati, proiezioni e statistiche. Dati e proiezioni formulati non di rado a lunghissimo termine, destinati ad essere smentiti dalla sopravvenienza degli inevitabili imprevisti.

La pervasività della rivoluzione digitale sviluppatasi negli ultimi anni forse troppo rapidamente implementa atteggiamenti regressivi di chiusura o comunque di diffidenza.

Per prima cosa, il linguaggio informatico si alimenta non di rado con "formule" arcane intrise di incomprensibilità comunicativa, con una metodologia di voluto involgimento che richiama il tradizionale uso del latino nelle celebrazioni religiose, di fatto estraneo alla moltitudine dei credenti.

Se realmente si vuole digitalizzare la società (anzi, il globo tutto), secondo un processo definito rivoluzionario, occorre quanto meno definire un sistema di richiamo comunicativo chiaro nella sua sinteticità innovativa ed evocativa.

Così la parola "algoritmo", misteriosa e arrogante insieme, potrebbe acquistare in simpatia ove ad esempio la si ricollegasse al suo curioso passato. La parola infatti è una traslitterazione del nome del matematico arabo *Alkhwārizmī*, che nel nono secolo contribuì con i suoi scritti, soprattutto il fondamentale *Sul calcolo con i numeri hindi*, al passaggio in Europa dalla numerazione romana a quella indo-araba. La parola algoritmo per assonanza è stata poi collegata al latino *algorismus*, assumendo il significato di metodo di calcolo, così invasivamente resuscitato qualche secolo dopo.

Da qui la necessità di costruire in tempi brevi percorsi culturali di indirizzo qualitativo di questa rivoluzione, che non può e non deve essere confinata nel solo momento tecnologico e informatico.

Infatti, la rivoluzione digitale non può rinchiudersi nella imperante e deleteria *datocrazia* che cerca di rendere numero il futuro e racchiudere i sogni in una curva statistica.

La storia – questa sconosciuta dei nostri tempi – ci insegna che l'umanità ha incontrato in varie epoche dei periodi caratterizzati dalla velocizzazione dell'innovazione delle conoscenze e della capacità di veicolarle su larga scala.

A titolo di esempio, basti ricordare gli ultimi periodi dell'Impero Romano, ai quali è mancata solo la scoperta dell'energia non umana, o anche i meravigliosi anni della *Bella Époque*.

In entrambi i riferimenti storici, alla grande crescita tecnologica sono purtroppo subentrato catene di eventi devastanti, quali il regresso medioevale (fra l'altro con l'iconizzazione della scrittura) e la Prima guerra mondiale.

Il primo obiettivo della rivoluzione digitale e della sua *governance* deve essere quello di porsi al servizio di percorsi conoscitivi eticamente virtuosi, ancorati a solide catene valoriali.

L'umanesimo digitale deve guidare i processi innovativi e le tecnologie di avanguardia, ricordando che l'uomo non può non restare al centro dei processi informatici, sia come loro "*dominus*" sostanziale, sia come destinatario reale delle utilità che indubbiamente si originano dalla espansione dell'intelligenza artificiale. La centralità dell'uomo deve essere rinforzata anche nelle cosiddette intelligenze artificiali generative, in grado di sviluppare, nella relativa ed apparente autonomia, capacità di ricerca delle fonti di conoscenza e di interazione con altre intelligenze. Autonomia solo apparente in quanto "a monte" è sempre presente l'addestramento umano.

L'umanesimo digitale – come ripetutamente sottolineato da Paolo Benanti in numerosi scritti (fra molti: *Le macchine sapienti. Intelligenze artificiali e decisioni umane* e *Noi e la macchina. Un'etica per l'era digitale*) – non deve tecnologizzare l'uomo, bensì umanizzare la macchina.

In questo modo il progressivo sviluppo ed il consolidamento del digitale potranno forse invertire la ricorrente tendenza "catastrofica postrivoluzionaria".

Si può sostenere dunque che la rivoluzione digitale avrà un senso se giungerà a consolidarsi in un ambito dove i processi informatici superino le connotazioni di esclusività autoreferenziale, purtroppo ad oggi molto diffuse.

(*) A. Avoli, Presidente onorario della Corte dei conti.

E la rivoluzione digitale avrà un senso solo se riuscirà ad attraversare le barriere del tecnicismo, che esclude larghi strati di popolazione, dando spazio ad una sorta di dittatura dei processi conoscitivi e ad una pericolosa compressione degli spazi della libertà individuale, non in grado di competere, e neppure di dialogare, con mondi e codici comunicativi troppo sofisticati e lontani dalla realtà fisica.

L'umanesimo digitale valorizza l'uomo nelle sue peculiarità specifiche, non clonabili neppure dalla più sofisticata delle intelligenze artificiali, in primo luogo l'immaginazione, la capacità di sognare e di interagire con spirito critico.

La tecnologia può ricostruire le emozioni, ma non potrà mai crearne o goderne. L'uomo affamato nella giungla davanti ad una banana lontana cento metri con un leone in agguato non deciderà se sfidare la sorte in base ad un calcolo informatico, ma solo in relazione al suo essere coraggioso, prudente o timoroso.

Ricondurre la rivoluzione digitale all'uomo non significa disconoscere i risultati di utilità concreta conseguiti applicando metodi di calcolo algoritmici, in grado di elaborare in tempi brevi quantità infinite di dati, secondo modelli addestrativi sempre più sofisticati,

La rivoluzione digitale ha avuto ed ha ricadute impressionanti in tutti i campi della vita quotidiana delle persone, dalla sanità, all'ambiente, dall'informazione alla agricoltura, dal mondo giuridico al restauro delle opere di arte.

Inoltre, ha conseguenze anche nel campo educativo e di trasmissione del sapere. Per la prima volta nella storia dell'umanità si è invertita la catena generazionale: oggi sono i nipoti che insegnano ai nonni ad acquisire una qualche idoneità informatica, ma i nonni non dovrebbero dismettere il loro ruolo di testimonianza dei valori e dell'indispensabile consapevolezza del significato del *tempo* come misura della vita.

I numerosi contributi presenti in questo quaderno si sono soffermati su molteplici aspetti, tutti però in qualche modo riconducibili alle tematiche generali appena evidenziate.

Più in particolare, il filo che accomuna i singoli contributi unisce le riflessioni sul quesito relativo a individuare e circoscrivere le modalità e i contenuti delle reazioni del mondo giuridico alle sfide dell'intelligenza artificiale e alle sue applicazioni algoritmiche.

Il diritto eurounitario e quello nazionale si sono dotati recentemente di specifici atti regolamentari che hanno cercato di porre rimedio ad un pericoloso vuoto di tutele, che la stessa giurisprudenza ha tentato a sua volta di coprire.

Nel campo dell'amministrazione pubblica l'informatizzazione dei procedimenti conduce ad una spersonalizzazione dei tradizionali centri di competenza, anche se il ruolo del *responsabile del procedimento* mantiene una posizione propria nel momento decisionale conclusivo, posizione che però nei fatti rischia di configurarsi più formale che sostanziale.

In linea astratta, le decisioni amministrative algoritmiche trovano spazio soprattutto nei procedimenti amministrativi vincolati, mentre recedono di fronte agli spazi di discrezionalità.

È emersa, in molti contributi, la preoccupazione per l'attenuazione delle garanzie, e quindi dei diritti, dei cittadini, protetti e tutelati costituzionalmente.

La soluzione regolamentare adottata si incentra sul principio di trasparenza delle fonti algoritmiche e dei parametri di addestramento dei processi elaborativi.

Una soluzione di principio assolutamente condivisibile, ma non sufficiente a colmare il *gap* di conoscenza tecnica, gap di sicuro aumentato con la informatizzazione delle procedure. Il cittadino medio difficilmente riesce a comprendere a pieno le ragioni per le quali, applicando determinati algoritmi, di per sé complessi, è maturata la decisione o, meglio, la bozza di decisione, dell'intelligenza artificiale. Può sembrare paradossale, ma il cittadino potrebbe essere costretto a ricorrere ad un algoritmo per contrastare i risultati di altro algoritmo.

Occorre avere il coraggio di ammettere che i risultati dei calcoli algoritmici utilizzati dall'intelligenza artificiale mantengano essi stessi il carattere della relatività, proprio di qualsiasi altra attività umana, influenzati come sono dagli *input* di addestramento. Tali *input*, umanissimi, non sfuggono ai rischi di strumentalizzazioni dolose e persino a impostazioni demagogiche volte ad alterare il libero gioco democratico. E neppure sfuggono all'errore propriamente detto.

Ad esempio, in una procedura di gara o di concorso assunzionale potrebbero essere inseriti in fase di addestramento elementi di pregiudizio discriminatorio, in ipotesi fondato su appartenenze religiose o razziali.

Oppure, nel diverso ambito del credito bancario, l'intelligenza artificiale può portare a qualificare una persona non affidabile, includendola nelle cosiddette centrali di rischio, senza possibilità di contraddittorio, anzi ad insaputa dell'interessato.

Se bene utilizzata, l'intelligenza artificiale costituisce un supporto alla fase decisionale, arricchendo i contenuti delle istruttorie in modo esponenzialmente positivo. Le preoccupazioni per l'eventualità di un uso distorto non debbono ostacolare l'applicazione, ma nondimeno non debbono venire meno adeguate procedure di controllo, interno e/o esterno imparziale, imperniate su indicatori parametrici di criticità, anche indirette, opportunamente preindividuate e anch'esse eventualmente supportate dall'informatica.

Gli operatori del Diritto qualche volta si arrovellano a lungo su questioni che sembrano realmente lontane da qualsiasi logica di buon senso.

In particolare, ci si è chiesti se alle intelligenze, soprattutto se generative, fosse possibile attribuire la personalità giuridica digitale, ovvero la titolarità di diritti, di doveri (*sic!*), di posizioni giuridiche protette.

La risposta è stata negativa e non poteva che essere tale: l'intelligenza artificiale è e deve restare uno strumento, sofisticato, complesso, meraviglioso, ma sempre strumento. Le sue *azioni*, in ipotesi dannose, vanno imputate agli addestratori o ai costruttori dei software (si pensi ai veicoli a guida autonoma).

Per altro verso, i brevetti ottenuti grazie all'intelligenza artificiale hanno sempre un'imputazione umana.

In sostanza il diritto nel suo complesso ha il merito di essersi, per primo, posto il problema di *calare nella realtà fisica* quella digitale, che deve dialogare con la prima senza annullarla, anzi, in qualche misura, porsi in suo subordine.

La negazione della personalità giuridica digitale, gli obblighi di trasparenza delle fonti algoritmiche e dei criteri di addestramento sono punti fermi in un panorama, come ricordato, di incontrollata pervasività informatica.

Le misure regolamentari europee e nazionali sono insufficienti ad eliminare le criticità della rivoluzione digitale. Tuttavia, hanno un valore di principio non secondario, ricollegandosi a valori insostituibili quali la dignità umana, la libertà individuale in tutte le sue declinazioni.

Nel campo processuale, si sono avviate, soprattutto negli Stati Uniti, esperienze di giurisprudenza cosiddetta predittiva. Il tema, intrigante di per sé, riveste un rilevante carattere *simbolico* rispetto alla definizione del rapporto fra la decisione umana e quella digitale.

La sentenza è atto tipico di un processo di ponderazione valutativa di interessi contrapposti, ponderazione ancorata ai correlati presupposti normativi, alla loro interpretazione e al quadro probatorio offerto dalle parti in giudizio.

La giustizia predittiva configura processi decisori automatici, nei quali confluiscono tutti i dati identificativi della fattispecie, mentre l'identificazione e l'applicazione delle regole di diritto avviene mediante una sorta di abbinamento automatico, supportato dai precedenti valorizzati in base al criterio dell'analogia.

Il sistema italiano non ammette la decisione predittiva, mantenendo fermo il principio della attribuzione al *giudice persona* della competenza e della responsabilità della sentenza.

Tuttavia, è proprio nel campo processuale che l'intelligenza artificiale si sta rivelando di particolare utilità, offrendo al giudicante una banca dati quanto mai articolata e in questo modo determinando indirettamente percorsi nomofilattici altamente attendibili.

Dunque, proprio dal mondo del diritto, sembrano emergere alcune delle chiavi di governo della rivoluzione digitale.

Come tutte le rivoluzioni, anche questa, dopo la fase eversiva, acquisterà una dimensione, per così dire, istituzionale. Anche questa rivoluzione diventerà, prima o poi, il passato, soggiacendo all'inevitabile *app* del tempo. E ad essa seguirà nella storia dell'umanità un'altra successiva rivoluzione. Con l'augurio che non si tratti di un ritorno alla vita preistorica in caverna.

* * *

Direzione e redazione

Via Antonio Baiamonti, 25 - 00195 Roma - tel. 0638762191 - E-mail: massimario.rivista@corteconti.it



ISBN 978-88-498-5743-6



I.S.S.N. 0392-5358
(c.m. 30-443900120101)