



ELIO FAMELI

in collaborazione con **ENRICO BISENZI**

La nuova Intelligenza artificiale e il problema della sua regolazione giuridica. La strategia europea e il lungo iter legislativo dell'AI Act

La rivoluzione tecnologica in corso in questi ultimi anni, soprattutto nel campo dell'Intelligenza artificiale (IA) e dell'apprendimento automatico, ha aperto scenari assolutamente nuovi e, fino a pochi anni fa, del tutto inimmaginabili. Il computer non è più la macchina inerte che si limita a eseguire le istruzioni impartite dal programmatore, ma, in quanto dotato di "capacità di apprendimento", diviene interlocutore attivo, in grado non solo di registrare enormi masse di dati, ma anche di elaborarle estraendone nuova conoscenza. Di fronte a questo rapido sviluppo tecnologico, l'Ue si è trovata nella necessità di affrontare con urgenza il problema della sua disciplina giuridica. La Legge europea sulla IA (AI Act), recentemente entrata in vigore, ponendosi a conclusione di un lungo e faticoso percorso, pur con i suoi inevitabili limiti, si configura come prima esperienza di regolazione organica e strutturata dell'Intelligenza artificiale nella molteplicità delle sue applicazioni.

*Intelligenza artificiale – Apprendimento automatico – Rapporto uomo-computer – Strategia europea per l'IA
AI Act – Considerazioni critiche sull'AI Act*

The new Artificial Intelligence and the problem of its legal regulation. The european strategy and the long legislative process of the AI Act

The technological revolution that has taken place in recent years, especially in the field of Artificial Intelligence (AI) and machine learning, has opened up completely new scenarios that were unimaginable just a few years ago. The computer is no longer an inert machine that simply executes the instructions given by the programmer in a rigid and sequential manner but, because it is endowed with "learning capabilities", it becomes an active interlocutor, capable not only of absorbing enormous amounts of data, but also of processing them, extracting new knowledge from them. In the face of this rapid technological development, the EU was faced with the urgent need to address the problem of its legal regulation. The European AI Act comes at the end of a long and arduous journey and, despite its inevitable limitations, is the first experience of organic and structured regulation of Artificial Intelligence in the multiplicity of its applications.

*Artificial Intelligence – Machine Learning – Human-computer interaction – European AI Strategy – AI Act
Critical remarks on the AI Act*

La Rivista è felice di ospitare le riflessioni proposte in questo saggio da Elio Fameli, componente del consiglio scientifico della RIID, in collaborazione con Enrico Bisenzi. E. Fameli è dirigente di ricerca in quiescenza del CNR (Istituto di Informatica Giuridica e Sistemi Giudiziari - IGSG di Firenze). E. Bisenzi è docente di "Inclusive Design" (Settore: "Graphic Design") presso l'Accademia di Belle Arti di Roma. Il presente contributo è frutto della piena condivisione di idee e opinioni tra gli Autori

SOMMARIO: 1. Introduzione: l'uomo e la macchina a confronto, tra prospettive di progressi fantascientifici e timori di scenari apocalittici. – 2. La rivoluzione dell'Intelligenza artificiale: il paradigma dei giochi strategici e il passaggio dall'approccio analitico a quello euristico nella soluzione dei problemi complessi. – 2.1. Il punto di svolta nell'evoluzione dell'Intelligenza artificiale: dalla potenza di calcolo alla capacità di apprendimento. – 2.2. L'apprendimento automatico e la pervasività delle nuove tecnologie. – 2.3. La "Teoria della mente" e la nuova configurazione del rapporto uomo-computer. – 3. Il problema della regolazione giuridica dell'Intelligenza artificiale: l'approccio europeo. – 3.1. La diffusione dell'Intelligenza artificiale: opportunità e rischi. – 3.2. Il contesto geopolitico e gli obiettivi della politica europea per l'IA. – 4. La Legge europea sull'Intelligenza artificiale (AI Act): entrata in vigore e ambito di applicazione. – 5. Considerazioni finali: ultimi sviluppi dell'Intelligenza artificiale e prime valutazioni dell'AI Act.

1. Introduzione: l'uomo e la macchina a confronto, tra prospettive di progressi fantascientifici e timori di scenari apocalittici

Dopo una fase di apparente letargo, scarsa visibilità e diffuso scetticismo, da alcuni anni l'Intelligenza artificiale (IA) sta conoscendo un periodo di rigoglioso sviluppo e sempre più vaste applicazioni, fino al punto di configurarsi come il settore di ricerca più avanzato e coinvolgente, quello – in ultima analisi – in cui si dovranno presumibilmente registrare le conquiste più affascinanti del futuro. La presenza pervasiva dell'IA è ormai una costante del nostro vivere quotidiano: tra le mura domestiche come in auto, al lavoro come nel divertimento, ci troviamo continuamente – ma spesso anche in maniera inconsapevole – a interagire con strumenti, dispositivi e, più di recente, "assistenti intelligenti". Ogni giorno, ormai senza rendercene conto, interagiamo con "macchine intelligenti" che indirizzano le nostre telefonate, approvano gli acquisti con carta di credito, interpretano gli esami medici, compongono musica e creano persino opere d'arte. Presto, inoltre, le automobili a guida autonoma percorreranno le strade di tutto il mondo. Negli anni Cinquanta del Novecento sognavamo di costruire intelligenze simili a quella

dell'uomo. Oggi abbiamo creato macchine che, elaborando grandi quantità di dati, riescono a riconoscere e a prendere decisioni. Non ci chiediamo più neppure chi ci stia educatamente rispondendo al telefono, con chi stiamo utilmente conversando sul computer, chi ci stia fornendo preziose indicazioni sul viaggio da intraprendere o lo svago da preferire. Intorno a noi si moltiplicano "soggetti" senza volto, che dimostrano però di conoscere bene il nostro "profilo", anche nei suoi aspetti più delicati e sensibili, dalla collocazione sociale al livello culturale, dalle convinzioni religiose all'orientamento politico.

Dei rischi connessi agli sviluppi e alla diffusione dell'IA si sono occupate menti eccelse e personalità illustri, nel campo della scienza, della finanza, ma anche dell'economia e della politica. Così, secondo il celebre cosmologo Stephen Hawking, l'umanità avrebbe ormai raggiunto un punto di non ritorno in quanto l'IA arriverà presto a un livello tale da superare di gran lunga la specie umana, configurandosi come una nuova forma di vita. Se, da una parte, è incontestabile che i progressi nel campo dell'IA possono portare benefici inimmaginabili alla società, è comunque altrettanto certo che risulta inevitabilmente molto sottile la linea di confine tra i vantaggi che si possono ricavare dalla tecnologia e il pericolo che questa ci sfugga di mano: "l'AI

potrebbe essere il più grande evento nella storia della nostra civiltà, oppure il peggiore”¹.

Un giorno l’intelligenza delle macchine supererà quella degli esseri umani? Siamo dunque a un passo dall’Apocalisse, circondati da supercomputer in grado di prendere il sopravvento sull’uomo? La loro crescita e il continuo incremento della loro potenza preannuncia forse l’avvento d’una nuova era, dominata da macchine capaci di svolgere compiti complessi molto meglio e più velocemente di noi? Già diversi anni or sono questi allarmanti interrogativi emergevano, ad esempio, negli scritti raccolti da Douglas Heaven, senior editor per l’IA presso la “MIT Technology Review”².

Elon Musk³, insieme a un gran numero di personalità del mondo scientifico ed esperti di robotica e Intelligenza artificiale, già nel 2017, con una lettera aperta⁴ indirizzata all’ONU e resa pubblica in occasione della Conferenza mondiale sull’IA (IJCAI 2017)⁵, aveva lanciato un accorato allarme contro lo sviluppo e l’uso di “sistemi letali autonomi”, inevitabilmente destinati a provocare il disastro totale di una terza guerra mondiale. Il documento, che vedeva per la prima volta uniti i rappresentanti delle società di AI e robotica di 26 Paesi, anche per il gran numero di firmatari evidenziava la gravità e l’urgenza della situazione. Ormai non si tratta più di uno scenario ipotetico, bensì di un problema reale e concreto, che richiede risposte rapide e adeguate: a differenza di altre potenziali manifestazioni dell’IA che appartengono ancora al mondo della fantascienza, attualmente i sistemi autonomi armati sono al primo posto nella classifica degli investimenti per lo sviluppo di nuove tecnologie.

Molto di recente, sui rischi di uno sviluppo incontrollato dell’Intelligenza artificiale si sono espressi anche i due vincitori del Premio Nobel per

la Fisica, John Hopfield, fisico statunitense attivo soprattutto nel campo della biofisica e della fisica statistica, e Geoffrey Hinton, psicologo e informatico britannico naturalizzato canadese, considerato uno tra i ricercatori più influenti nello sviluppo dell’apprendimento profondo. Nella motivazione del premio è messa in evidenza la grande importanza delle ricerche svolte da questi due scienziati, che hanno utilizzato gli strumenti della Fisica per porre le basi dello sviluppo attuale dell’apprendimento automatico. La capacità d’imparare in modo autonomo, basata sulle reti neurali artificiali, sta rivoluzionando la scienza, l’ingegneria e persino la nostra vita quotidiana. Hinton, in particolare, ritiene che l’Intelligenza artificiale “avrà sull’umanità conseguenze paragonabili alla rivoluzione industriale: all’epoca le macchine ci superavano in forza fisica, ora invece sono destinate a superarci dal punto di vista intellettuale. Ci saranno effetti estremamente positivi, ... ma potranno esserci anche conseguenze negative, qualora le macchine riescano a sfuggire al nostro controllo”⁶.

Sul fronte opposto a quello dei “preoccupati” per il continuo potenziamento delle capacità dell’IA e per il suo inarrestabile dilagare in ogni settore dell’attività umana, non mancano studiosi e divulgatori che tendono a evidenziare i limiti e gli insuccessi di computer superpotenti o di robot umanoidi impegnati in compiti anche modesti⁷. Esistono corpose raccolte di storielle da raccontare sui fallimenti di robot pur sofisticati ed evoluti, ma sembrano comunque ingenuamente rassicuranti affermazioni categoriche come quella enunciata da Will Douglas Heaven⁸ in un suo articolo sull’argomento: “Far from being a steady march to greatness, the past and present of robotics and AI are littered with examples of banal practicalities tying machines down. If you want to look at what the

1. La citazione è in INTINI 2017.

2. Vedi NEW SCIENTIST 2018.

3. Elon Musk, imprenditore sudafricano naturalizzato statunitense, è cofondatore di OpenAI, un laboratorio di ricerca sull’Intelligenza artificiale costituito con lo specifico obiettivo di promuovere e sviluppare una IA amichevole, in grado di contribuire al benessere dell’umanità.

4. Si veda MUSK et al. 2017.

5. [IJCAI-17](#), Melbourne, August 19-25, 2017.

6. Cfr., *inter alia*, INTINI 2024; DUSI 2024; ROVELLI-OTTOLINA 2024.

7. Si vedano, tra i molti, ZREICK 2018; FLEISCHNER 2023.

8. [Will Douglas Heaven](#) è “senior editor” per l’Intelligenza artificiale presso la “MIT Technology Review”.

future of AI really holds, it's not the highlight reels that matter – it's the out-takes"⁹.

In realtà non mancano fatti imbarazzanti e sicuramente negativi che, a ben guardare, piuttosto che evidenziare i limiti dell'IA e delle sue applicazioni, mostrano i rischi che derivano dall'uso distorto che se ne può fare. Esempio, in questo senso, è stata la vicenda di Tay, il chatbot di Microsoft che aveva debuttato nel 2016 su Twitter: progettato per imparare autonomamente parole ed espressioni in linguaggio naturale dialogando direttamente con gli utenti sui social, dopo poche ore dall'attivazione aveva cominciato a scrivere commenti offensivi, sessisti e razzisti, costringendo così la Microsoft a porre fine all'esperimento nel giro di 24 ore¹⁰.

Per concludere questi brevi cenni sui rischi connessi allo sviluppo incontrollato dell'IA e sui limiti che – almeno per ora – sembrano contrassegnarne l'evoluzione, si può qui fare riferimento all'osservazione di Mark Riedl¹¹, secondo cui resteranno comunque fuori dalla portata del computer i linguaggi altamente veloci e complessi, con forti riferimenti contestuali e situazionali, ma anche l'umorismo, la comicità originale e innovativa, così come, per altro verso, le decisioni imprevedibili e le idee rivoluzionarie. In questo senso l'IA, basandosi sull'accesso a immense quantità di dati esistenti che poi sintetizza ed elabora, sarebbe una "tecnologia conservatrice". Com'è stato autorevolmente affermato: "Il computer, non essendo in grado di capire cosa siano i tabù, non può infrangerli"¹².

A queste ultime considerazioni, a un livello più generale e fortemente interdisciplinare, possono infine collegarsi le teorie rassicuranti di chi riafferma l'indefettibile superiorità dell'uomo

sulla macchina fondando la sua ricerca non solo sul rigore scientifico, ma anche sulla visionarietà tecnologica e sull'elaborazione filosofico-religiosa¹³. Partendo dalla constatazione della deriva materialista e riduzionista che sta portando a considerare l'uomo come una macchina, si evidenzia la necessità di una nuova scienza che includa la spiritualità e di una nuova spiritualità che, a sua volta, includa la scienza; si arriva così a proporre una rivoluzionaria, ma affascinante, teoria della realtà, in cui potrebbero trovare adeguata risposta i problemi della coscienza e del libero arbitrio¹⁴.

2. La rivoluzione dell'Intelligenza artificiale: il paradigma dei giochi strategici e il passaggio dall'approccio analitico a quello euristico nella soluzione dei problemi complessi

La trasformazione radicale cui è andata incontro l'Intelligenza artificiale nell'ultimo ventennio è espressa in maniera emblematica nell'evoluzione subita dall'approccio informatico ai giochi strategici, in particolare il gioco degli scacchi e quello del Go.

Rievocando gli entusiasmi, ma anche le paure e le polemiche, che nel 1996 e nel 1997 accompagnarono la storica gara tra Deep Blue, il calcolatore appositamente progettato dall'IBM, e Garry Kasparov, campione mondiale di scacchi, la vittoria schiacciante sull'uomo da parte dello stesso supercomputer nella versione riveduta e aggiornata, proprio per la filosofia che l'ha resa possibile, viene a configurarsi in maniera emblematica nei confronti delle radicali trasformazioni che negli ultimi due decenni si sono verificate nel campo dell'Intelligenza artificiale¹⁵.

9. Cfr. DOUGLAS HEAVEN 2017. Nello stesso senso si veda LANDGREBE–SMITH 2022.

10. Si vedano, tra gli altri, WOLF–MILLER–GRODZINSKY 2017; METZ 2018.

11. Mark Riedl, docente di "Interactive Computing" presso il "Georgia Institute of Technology" e direttore associato del "Machine Learning Center", conduce ricerche orientate allo sviluppo di tecnologie di Intelligenza artificiale e di apprendimento automatico in grado d'interagire in modo naturale con gli utenti umani.

12. L'osservazione è di Guy Hoffman (Cornell University) ed è riportata in BAUM 2023.

13. Tra le opere più recenti e autorevoli, si vedano, in particolare, FAGGIN 2022.

14. In questo senso FAGGIN 2024.

15. Sulla sfida tra il supercomputer dell'IBM e il russo Garry Kasparov, campione del mondo di scacchi per ben 15 anni (1985-2000), cfr. HSU 2004. Alcune riflessioni sul confronto uomo-macchina sono anche in ZUCCONI 1997.

Ugualmente, la sfida a Go¹⁶ tra il software AlphaGo, sviluppato dalla società DeepMind di Google, e il sudcoreano Lee SeDol, uno dei campioni mondiali di questa specialità, si è conclusa nel marzo del 2016 con una sonora sconfitta per l'uomo nei confronti della macchina: un solo match vinto dall'uomo contro quattro vinti dalla macchina. Quello che è stato definito uno degli eventi più importanti nella storia dell'Intelligenza artificiale aveva avuto un precedente, con la sconfitta del campione europeo Fan Hui nel gennaio del 2016. Nel maggio 2017, poi, si era avuta un'altra conferma delle capacità, non solo logiche ma in qualche modo anche intuitive, del supercomputer di Google, con la vittoria sul cinese Ke Jie, detentore del titolo di campione del mondo di questo millenario gioco strategico.

La complessità dei due giochi è estrema tanto che, superati certi parametri, può essere soltanto immaginata. Negli scacchi il numero delle combinazioni ammesse dei 32 pezzi sulle 64 caselle della scacchiera si calcola sia compreso fra 1043 e 1050, mentre la dimensione del c.d. "albero" delle mosse sarebbe addirittura di 10123. Nel Go, su una scacchiera (*goban*) con una griglia da 19 x 19 caselle i due avversari collocano pedine ("pietre") nere o bianche con l'obiettivo di controllare la zona più ampia della superficie di gioco, ma il numero limitato e la semplicità delle regole nascondono una quantità di combinazioni possibili largamente superiore a quella calcolata per gli scacchi¹⁷.

Il passaggio dalla sfida tra Deep Blue e Garry Kasparov (1996-1997) a quella tra AlphaGo e Lee SeDol, Fan Hui e Ke Jie (2016-2017) può essere assunto a rappresentazione emblematica della rivoluzione in atto nel mondo dell'Intelligenza artificiale. Nel primo caso lo sforzo dei progettisti informatici si era concentrato sulla potenza di calcolo. Deep Blue, nella versione perfezionata nel 1997, non era un normale calcolatore elettronico, bensì un supercomputer ad altissimo parallelismo, con una rete di calcolo a 32 nodi, dove ogni singolo nodo utilizzava una scheda microchannel a 8 microprocessori VLSI (*Very Large Scale Integration*), per un totale di ben 256 microprocessori. Con questa struttura la macchina era in grado di calcolare, in condizioni ottimali, più di un miliardo di mosse al secondo e 60 bilioni in 3 minuti¹⁸. Diversamente da Deep Blue, il software utilizzato da AlphaGo per battere ripetutamente – e definitivamente – i campioni del mondo di questo difficilissimo gioco strategico è, invece, essenzialmente costituito da una combinazione tra tecniche d'apprendimento automatico (*Machine Learning*), basate sull'impiego di "reti neurali profonde", e tecniche di ricerca ad albero. A queste, però, è affiancata una fase estensiva d'apprendimento da gioco, in cui la macchina è messa in grado di memorizzare ed elaborare centinaia di migliaia di partite giocate da campioni umani o da altre macchine adeguatamente addestrate. I dati in input, prima di essere inviati alle reti neurali, nella fase di *pre-processing*, vengono analizzati in modo da individuarne le

16. Il Go è un gioco da tavolo strategico che ebbe origine in Cina, dove è giocato da almeno 2.500 anni. Divenuto rapidamente molto popolare nell'Asia orientale, dalla fine del XIX secolo ha cominciato a diffondersi anche in Occidente. Una presentazione sintetica, ma articolata di questo gioco, della sua storia e delle sue regole, ma anche della sua rilevanza nella letteratura, nella musica e nel cinema, oltre che nella matematica e nell'informatica, è in *Go (gioco)*, Voce in "Wikipedia. L'enciclopedia libera". Recentemente, della sfida a Go tra il software AlphaGo di Google e il campione sudcoreano Lee SeDol è tornata a occuparsi anche la stampa italiana. Cfr. LA REPUBBLICA (Red.) 2025.

17. Un proverbio coreano dice che nessuna partita di Go è mai stata giocata due volte, affermazione che risulta verosimile. Solo agli inizi del 2016, grazie all'impiego di software sofisticati e di un hardware particolarmente potente, siamo stati in grado di calcolare il numero di posizioni che le pedine possono assumere nel gioco. Le combinazioni ammesse sarebbero pari a 2,08 per dieci elevato alla centosettantesima: si tratta non di un numero infinito – come molti avevano sostenuto per secoli –, ma comunque di una cifra superiore al numero totale di atomi osservabili nell'universo. Il calcolo esatto delle posizioni ammesse per le pedine nel Go si deve a due informatici: John Tromp e Michal Koucký. Per considerazioni più approfondite si rimanda a *Go and Mathematics. Legal Positions*, Voce in "Wikipedia. L'enciclopedia libera".

18. Per una sintetica, ma chiara descrizione del supercomputer realizzato dall'IBM per il gioco degli scacchi si veda GLUSHENKOV 2024.

principali caratteristiche e, in particolare, l'eventuale corrispondenza a modelli ricorrenti (*pattern*). Le reti neurali sono state inizialmente modellate sulla base del comportamento dei giocatori umani: con l'obiettivo d'imitarlo è stato utilizzato un database in cui sono stati memorizzati milioni di mosse tratte da partite storiche¹⁹.

2.1. Il punto di svolta nell'evoluzione dell'Intelligenza artificiale: dalla potenza di calcolo alla capacità di apprendimento

Evitando in questa sede d'addentrarsi in descrizioni tecniche approfondite, importa qui rilevare come la fondamentale differenza tra l'approccio informatico impiegato dai progettisti del supercomputer IBM, da una parte, e la filosofia dei costruttori dell'algoritmo di AlphaGo, dall'altra, risieda proprio nel passaggio dal ricorso alla "forza bruta"²⁰ della potenza di calcolo alla (ri)scoperta del ruolo imprescindibile della "capacità d'apprendimento", basata sull'esperienza²¹.

In questo senso l'importanza della vittoria del programma di Google sul campione mondiale in carica può ritenersi di portata storica, non solo con riferimento all'evoluzione delle tecniche di programmazione nel campo dell'Intelligenza artificiale, ma addirittura come pietra miliare nel processo di sviluppo scientifico e tecnologico dell'umanità. Alla precisione – ma anche alla rigidità – del calcolo logico-matematico vengono ad affiancarsi la flessibilità dell'adattamento, la prontezza della reazione, la velocità e l'efficienza dell'intuizione; in maniera fino a poco tempo fa del tutto

imprevedibile e, comunque, in un contesto di allarmata attenzione, si cominciano a delineare, per il computer adeguatamente programmato e "addestrato", potenzialità che si potrebbero far rientrare in categorie concettuali quali l'istinto, l'intuizione e persino la creatività.

Nella sfida tra l'uomo e la macchina i tempi si accorciano²², mentre la lotta diviene serrata: una ad una sembrano crollare le roccaforti degli "esseri umani", aggrappati a quelle prerogative essenziali (l'intelligenza, la volontà, il libero arbitrio, la responsabilità morale) che una riflessione filosofica millenaria ha individuato ed elaborato, ma anche costretti alla strenua difesa di una "superiorità" sempre più insidiata dagli "agenti artificiali". Alcune tra le tradizionali certezze cominciano pericolosamente a vacillare: l'immagine tranquillizzante del computer che esegue esclusivamente e fedelmente gli ordini (le istruzioni) che l'uomo (l'informatico programmatore) ad esso impartisce via via assume contorni meno netti, si affievolisce e s'incrina. All'orizzonte si affaccia – all'inizio timidamente, lontano dagli ambienti e dai modelli accademici, sovente nel chiuso dei laboratori informatici privati – il computer che "impara (da solo) dai dati" e che, sulla base dell'esperienza acquisita, "prende decisioni". Com'è evidente, si tratta di una vera e propria rivoluzione, capace d'incidere profondamente non solo nel mondo dell'informatica e delle sue sempre più vaste applicazioni, ma in generale in ogni settore della scienza e della cultura.

Questa nuova configurazione del ruolo del computer, proprio perché fuoriesce completamente dallo schema del mero esecutore di ordini, del

19. Dopo l'ultima vittoria contro il campione del mondo Ke Jie, con un comunicato stampa della DeepMind fu annunciato il ritiro di AlphaGo dalle competizioni. Gli scienziati che avevano lavorato alla sua realizzazione tennero a evidenziare che l'esperienza acquisita sarebbe stata sicuramente utilizzabile in altri settori applicativi, tra cui la "polizia predittiva" (*predictive policing*), la sanità e le energie rinnovabili. Si veda SILVER–HUANG–MADDISON et al. 2016.

20. Il metodo "Brute Force" ("forza bruta" o "ricerca esaustiva") in Informatica è un algoritmo di risoluzione di un dato problema, che consiste nel verificare tutte le soluzioni teoricamente possibili fino a che si trovi quella effettivamente corretta.

21. Si veda in MCAFEE–BRYNJOLFSSON 2016: "...So, why is AlphaGo's victory significant?...What changed? The AlphaGo victories vividly illustrate the power of a new approach in which instead of trying to program smart strategies into a computer, we instead build systems that can learn winning strategies almost entirely on their own, by seeing examples of successes and failures".

22. La vittoria di AlphaGo su Lee SeDol, il 15 marzo 2016, è arrivata improvvisa e inaspettata: qualunque esperto stimava che il dominio umano nel gioco del Go sarebbe rimasto incontrastato per un periodo di almeno 5-10 anni. Si veda MAGGIOLO 2016.

tutto controllabile e prevedibile, suscita immediatamente interrogativi e perplessità. Tornando all'emblematica sfida tra uomo e computer al gioco del Go, si è opportunamente messo in rilievo che la schiacciante vittoria conseguita dalla macchina sull'uomo non è stata – e non poteva essere – in alcun modo il risultato dell'analisi di tutte le combinazioni possibili nel posizionamento delle pedine. Come già precisato²³, il numero delle “posizioni ammesse” (*legal positions*) per i pezzi del gioco – numero che per secoli si era ritenuto fosse addirittura infinito – è stato calcolato solo molto di recente; si tratta, però, di un valore comunque non dominabile non solo dall'uomo, ma neppure utilizzando l'enorme potenza di calcolo dei computer oggi disponibili. Proprio per questo nessuno è in grado di spiegare come si giochi a Go ai livelli più alti: gli stessi giocatori più abili non riescono ad accedere alle conoscenze che permettono di sviluppare strategie di enorme difficoltà concettuale.

Il fenomeno è ampiamente conosciuto, perché in realtà non riguarda soltanto il gioco del Go, bensì molti tra i problemi che ci si trova quotidianamente ad affrontare nella vita quotidiana, dalla guida di un veicolo, in cui è necessario porre in essere sequenze di movimenti rapidi e coordinati, al riconoscimento di caratteri e di volti all'interno d'insiemi costituiti da migliaia di altri elementi simili. Come lo scienziato e filosofo ungherese Michael Polanyi²⁴ aveva efficacemente espresso, “sappiamo più cose di quelle che riusciamo a spiegare”: la vera conoscenza, dunque, non può essere integralmente formalizzata in regole o algoritmi, essendo soprattutto “conoscenza tacita”. Questa osservazione assume un'importanza cruciale nel dibattito sulla possibilità di costruire “macchine intelligenti”, in grado di risolvere problemi complessi, dal momento che consente di prescindere dalla necessità di possedere una conoscenza analitica, esplicita e organizzata della materia da trattare. Il celebre paradosso espresso nel dialogo platonico *Menone*, in cui Socrate stesso argomenta che all'uomo non

è possibile né cercare quello che sa (in quanto da lui già conosciuto), né cercare quello che non sa (perché non saprebbe cosa cercare), attraverso la rivisitazione speculativa di Polanyi, torna ad essere incredibilmente attuale. Di fronte a problemi con un grado di complessità molto elevato l'uomo è in grado di trovare soluzioni sulla base di conoscenze che egli stesso non è in grado di spiegare.

Per la realizzazione di un software in grado d'affrontare problemi di elevata complessità è necessario abbandonare la strada della previsione analitica di tutte le strategie alternative, affidando invece direttamente alla macchina il compito d'individuare le soluzioni volta a volta adeguate rispetto sia al contesto complessivo che alla sequenza delle singole situazioni critiche da superare. In questo nuovo tipo d'approccio la conoscenza che il sistema utilizza è basata sull'esperienza e, quindi, sulla valutazione di un numero altissimo d'esempi, sia di successi sia di fallimenti.

L'esplosione combinatoria – presente nel gioco degli scacchi, ma in misura molto superiore in quello del Go, come anche in quasi tutte le situazioni della vita reale in cui si ponga la necessità d'assumere decisioni in tempi ragionevoli – non consente d'avventurarsi nell'analisi d'un numero imprecisabile di possibili opzioni. L'impossibilità di prevedere e calcolare la successione e le modalità nei cambiamenti di stato rispetto alla situazione di partenza impone d'abbandonare la strada del calcolo come strumento necessario a valutare miliardi di combinazioni: anche nella pratica, di fronte a un problema da risolvere, non si esaminano tutte le possibili alternative e tutte le conseguenze di ciascuna alternativa (fatta eccezione per le ipotesi in cui esse siano poche o siano comunque riconducibili a un numero limitato). Quando la situazione è complessa si rende necessario restringere lo spazio virtuale in cui effettuare la ricerca della soluzione del problema e quindi “potare” l'albero dei percorsi astrattamente possibili per ottenere il risultato desiderato. Diventa possibile gestire l'esplosione

23. Si vedano le indicazioni fornite nella nota n. 16.

24. Michael Polanyi è stato uno dei maggiori esponenti della “nuova filosofia della scienza” nella seconda metà del ventesimo secolo. Insieme a Thomas Khun, in aperta polemica con Karl Popper, fu protagonista di quella stagione cruciale del pensiero filosofico che portò a una profonda revisione e, quindi, al rinnovamento del modo d'intendere e valutare la ricerca scientifica. Il concetto di “conoscenza inespressa” o “conoscenza tacita” (*tacit knowledge*) è sviluppato nelle opere più note di questo autore, peraltro ancora poco conosciuto e studiato. Si vedano: POLANYI 1958; POLANYI 1966.

combinatoria solo avvalendosi di una “euristica” che consenta di scartare rapidamente alcune tra le opzioni teoricamente configurabili in modo da concentrare l'attenzione su un numero limitato di alternative rimaste.

2.2. L'apprendimento automatico e la pervasività delle nuove tecnologie

In via preliminare occorre, intanto, prendere coscienza del fatto che anche l'utente medio del computer ottiene informazioni e utilizza servizi che sono comunque il risultato di forme più o meno complesse di “apprendimento automatico” da parte della macchina. Quando si effettua una ricerca nella Rete, i dati che si ottengono in risposta – come anche gli annunci pubblicitari che ne rendono più lenta e faticosa la lettura – sono stati previamente selezionati e organizzati da complessi algoritmi di *machine learning*. L'utilizzo dei servizi di *social networking*, ormai sempre più diffuso e coinvolgente soprattutto su piattaforme come Facebook e X, ma anche le varie possibilità d'acquisto e vendita di beni e servizi nel Web si avvalgono ampiamente delle tecniche di autoapprendimento. Ugualmente avviene per la segnalazione di un libro come per l'offerta di un “pacchetto vacanze”, per le proposte di brani musicali o la scelta di eventi d'intrattenimento, tutte in qualche modo “calibrate” sulle preferenze dell'utente.

La presenza attiva del *machine learning*, tuttavia, non è limitata al cyberspazio: nell'arco della giornata ci accompagna, invisibile e discreta, ottimizzando i consumi del veicolo che guidiamo, controllando e adattando in continuazione la temperatura degli

ambienti in cui viviamo, agevolando e supportando il lavoro d'ufficio con “software intelligenti”, in grado di selezionare e organizzare la posta elettronica, estrarre le informazioni rilevanti da grandi masse di dati, effettuare traduzioni e controllare la correttezza ortografica e grammaticale dei testi. Con l'aiuto di “sistemi di apprendimento automatico” possono essere gestiti i nostri risparmi dalle banche, può essere approvata la richiesta di un mutuo, può essere controllato l'uso di una carta di credito, eseguita l'analisi dei test per l'ammissione a un corso di studi, effettuata la valutazione dell'immobile che si vuole vendere o comprare, persino decisa la dislocazione delle pattuglie di polizia nelle aree urbane a maggior rischio di criminalità. Tutto questo, e molto altro ancora, nel mondo che ci circonda come all'interno della nostra sfera privata, già oggi, sovente a nostra insaputa, costituisce il risultato delle complesse elaborazioni svolte dagli “algoritmi di apprendimento”.

Una rivoluzione profonda sta trasformando la nostra vita, a livello sia individuale che sociale, ma anche la scienza, la tecnologia, l'economia, la finanza, la politica e la guerra stessa. Il *machine learning* viene a configurarsi come l'unico strumento potenzialmente in grado di gestire e convertire in conoscenza le enormi masse di dati che ci vengono fornite dalle moderne forme d'indagine della natura e dell'uomo. Tra queste ci si limita qui a citare i satelliti artificiali²⁵, gli acceleratori di particelle²⁶ e i sequenziatori di DNA²⁷.

L'economia, la finanza, il mondo degli affari si avvalgono ampiamente del *machine learning*. Mai prima d'ora era stato possibile ottenere una

25. Con il termine “satellite artificiale” ci si riferisce a tutti gli oggetti orbitanti intorno a un corpo celeste, posti volutamente nell'orbita desiderata con mezzi tecnologici, con varie finalità di servizio, d'indagine o di monitoraggio scientifico-ambientale. L'insieme di più satelliti artificiali adibiti a uno stesso scopo forma una “costellazione” o “flotta”. Da quando, nel 1957, l'allora Unione Sovietica lanciò lo Sputnik 1, si calcola che siano stati messi in orbita intorno alla Terra oltre 14.000 satelliti artificiali; di questi, però, oltre il 75% non sarebbe più in funzione, affollando pericolosamente lo spazio intorno al nostro pianeta.

26. Gli acceleratori di particelle sono macchine tecnologicamente complesse che accelerano particelle microscopiche tramite l'interazione della carica elettrica con campi elettromagnetici statici e dinamici; è la realizzazione tecnica di questa interazione che differenzia i vari tipi di acceleratori. Le applicazioni sono divenute sempre più numerose: all'inizio gli acceleratori sono stati sviluppati essenzialmente per studiare gli elementi costituenti della materia e le leggi che regolano la loro interazione; in seguito, in particolare i collisori particella-antiparticella si sono rivelati utili sia per applicazioni tecnologiche che per applicazioni mediche.

27. Il genoma umano comprende centinaia di milioni di frammenti di codice genetico. Una mole di dati enorme, impossibile da analizzare in tempi brevi con l'ausilio della sola mente umana. Il sequenziamento del DNA consente ai ricercatori di decodificare tutti questi elementi, aiutandoli a svelare eventuali informazioni genetiche

conoscenza tanto approfondita, dinamica e continuamente aggiornabile, dei clienti, dei loro orientamenti nei consumi, dei loro gusti e, più in generale, delle loro stesse abitudini di vita. Nei negozi delle grandi catene di distribuzione gli approvvigionamenti delle merci e la loro stessa disposizione negli scaffali per la vendita avvengono sulla base di “modelli” che, a loro volta, governano “algoritmi di apprendimento”. Così pure accade in un altro importante settore economico, quello della scelta dei libri verso cui far convergere gli acquisti dei lettori, all’interno di quel fenomeno sempre più vasto e coinvolgente che è il commercio elettronico.

Nel campo della finanza può essere qui sufficiente richiamare l’evidente incidenza che in essa hanno – a livello sia nazionale che transnazionale – i modelli per la scelta dei titoli azionari da vendere e/o acquistare. La maggior parte dei fondi d’investimento sceglie i titoli su cui investire avvalendosi di algoritmi di apprendimento, ma non mancano neppure attività interamente gestite da sistemi di apprendimento automatico. Si può ragionevolmente presumere che – in un contesto in cui i mercati finanziari non sono più luoghi fisici, ma piattaforme informatiche (“sedi di negoziazione”) dove s’incrociano le proposte di acquisto e di vendita – i modelli per la selezione degli strumenti finanziari in realtà governino l’andamento della finanza mondiale in maniera meno trasparente, ma sicuramente più incisiva di qualunque autorità formalmente preposta alla regolazione e al controllo del settore.

Nei rapporti delle banche con i loro clienti l’impiego degli algoritmi di apprendimento assume sovente un ruolo determinante. Sostituiti programmi d’Intelligenza artificiale – solitamente sviluppati con caratteristiche funzionali e secondo parametri quasi del tutto privi di trasparenza per l’utenza – sovrintendono a numerose procedure operative: dall’apertura del conto corrente all’esame dei requisiti per il rilascio della carta di credito,

dal controllo delle eventuali transazioni sospette alla valutazione delle richieste di prestiti e mutui, fino alla rilevazione e all’analisi della c.d. *customer satisfaction*, in cui le aspettative dell’utenza rispetto a una determinata tipologia di prodotto o servizio sono messe a confronto con le percezioni/valutazioni delle corrispondenti esperienze di uso o consumo.

Ugualmente, nell’ambito della tutela dell’ordine pubblico e della sicurezza dei cittadini ci si avvale in vario modo del *machine learning*. Per esempio: sulla base del controllo delle immagini registrate dalle telecamere di sorveglianza possono essere segnalate attività sospette alla centrale di coordinamento; sofisticati algoritmi di previsione statistica sono in grado di suggerire agli organi di polizia la più efficiente dislocazione delle pattuglie di ronda; complessi software d’identificazione biometrica consentono di raccogliere dati personali confrontandoli in tempi rapidissimi all’interno di vasti archivi specialistici. Nella lotta al crimine già si configurano scenari fantascientifici in cui unità di pubblica sicurezza formate da automobili senza guidatore pattuglieranno le strade delle metropoli in stretto contatto con le centrali operative, in modo da innalzare il livello d’efficienza nella ricerca e nella segnalazione di persone, cose o situazioni sospette.

In generale, poi, nella vita quotidiana ci si avvale continuamente – e talora anche inconsapevolmente – di strumenti tecnologici e dispositivi elettronici (computer, tablet, smartphone) fortemente caratterizzati, nello svolgimento stesso delle loro funzioni, dal ricorso a meccanismi di auto-apprendimento. Veicoli senza conducente attraversano ormai in gran numero la terra, il mare e il cielo, pilotandosi da soli, né appartengono più alla fantascienza i progetti per la realizzazione di robot-postino, robot per lo svolgimento di lavori domestici²⁸, robot umanoidi²⁹ o droni per la consegna a domicilio di libri o alimenti. I nostri cellulari,

nascoste. In questo modo è possibile accedere a informazioni che riguardano il funzionamento del corpo umano, la crescita, l’invecchiamento e anche l’insorgere di malattie.

28. I robot domestici (*home robot*), ormai sempre più sofisticati e versatili, sono in grado non solo di cucinare e pulire, ma anche di sorvegliare la casa e controllare gli animali domestici. Anche in Europa progetti di robot umanoidi, in grado d’assistere gli anziani e le persone con ridotte capacità motorie, coinvolgono già da alcuni anni aziende leader e laboratori di ricerca impegnati nel campo dell’Intelligenza artificiale.

29. La “robotica umanoide”, orientata alla costruzione di “macchine intelligenti antropomorfe” in grado d’interagire con l’uomo e adattarsi all’ambiente circostante, è ormai definitivamente uscita dalla fantascienza per diventare

compagni ormai inseparabili di lavoro, di viaggio e di svago, sono molto più che semplici strumenti di comunicazione: trasmettono, ricevono e organizzano – anche autonomamente – documenti e immagini, correggono gli errori di battitura, suggeriscono e completano le parole che inseriamo nei nostri testi, interpretano i comandi vocali con cui possiamo formulare richieste al nostro “Personal Assistant”. Agli Assistenti Personali³⁰ installati sugli smartphone è possibile rivolgersi, in lingua naturale, per ottenere rapidamente informazioni che spaziano dalle previsioni meteorologiche alla ricerca di luoghi d’interesse, dalle indicazioni stradali e geografiche alle classifiche sportive, dalla descrizione di luoghi e persone alla predisposizione di promemoria e alla effettuazione di calcoli matematici e conversioni valutarie.

In ultima analisi, però, se, da una parte, si ottengono dal computer le risposte desiderate in corrispondenza alle domande che a esso sono state rivolte, dall’altra, il software impiegato acquisisce gradatamente una conoscenza sempre più precisa e circostanziata di chi lo utilizza in maniera regolare e continuativa, dei suoi dati personali, delle sue capacità lavorative e competenze professionali, ma anche delle sue abitudini di vita, delle sue relazioni sociali, come pure dei suoi gusti e orientamenti di consumo. Di tutta questa enorme massa di dati, di questa scia invisibile che ogni utente del Web, più o meno inconsapevolmente, lascia dietro di sé dopo ogni consultazione, si alimentano motori di

ricerca sempre più potenti e sofisticati. Ogni interazione col computer (nelle sue varie declinazioni più recenti, che attualmente possono estendersi dal grosso server dislocato a migliaia di chilometri di distanza al piccolo smartphone d’uso quotidiano) in effetti avviene su due livelli distinti e in due direzioni opposte: nell’immediato l’utente consegue il risultato cui tende, sia esso la risposta a una domanda, la vendita o l’acquisto di un prodotto, l’esito di una transazione commerciale o di un’operazione finanziaria; sotto un diverso profilo, però, è invece il computer a ottenere qualcosa dall’utente, acquisendo i suoi dati personali, insieme a una serie d’informazioni precise e in vario modo articolate sulla sua vita e i suoi interessi.

Occorre, inoltre, osservare che, tralasciando i dati che vengono comunicati al computer in maniera trasparente e consapevole (quali i dati identificativi, anagrafici, telefonici, sanitari, gli indirizzi e-mail e i vari account richiesti dai motori di ricerca e dai servizi di *social networking*), molte altre informazioni sull’utente sono acquisite in forme sovente occulte e insidiose: per esempio, attraverso testi, documenti, registrazioni di percorsi di ricerca sulla Rete, rilevamenti effettuati dai microprocessori degli autoveicoli e dai sistemi di geolocalizzazione dei cellulari, immagini e sequenze riprese da telecamere di sorveglianza. Una così grande quantità e varietà di dati, poi, non rimane inerte e inutilizzabile, dal momento che sono

un campo di ricerca emergente, destinato ad attirare sempre più l’attenzione di ricercatori e case costruttrici. Robot simili all’uomo – dotati di braccia, gambe, sensori tattili e visivi – erano stati preconizzati dal drammaturgo ceco Karel Čapek, che per primo usò questo termine nel suo dramma dal titolo *R.U.R. Rossumovi univerzální roboti* (“I robot universali di Rossum”), pubblicato nel 1920. Nei prossimi decenni questi androidi programmabili potranno diventare nostri compagni affidabili e servizievoli, svolgendo compiti ripetitivi, sgradevoli o altamente rischiosi. Già oggi, comunque, le realizzazioni sono numerose e interessanti, spaziando in settori applicativi diversi, dalla medicina all’assistenza agli anziani, dall’intrattenimento alle operazioni di soccorso in zone di guerra o colpite da calamità naturali. Robot umanoidi sono stati costruiti soprattutto in Giappone e in Corea, ma sperimentazioni rilevanti sono in corso anche in Europa. R1 è il nome del robot progettato e realizzato in Italia dall’Istituto Italiano di Tecnologia (IIT) di Genova; le sue capacità d’apprendimento automatico lo rendono adatto a svolgere compiti di accoglienza, assistenza e sorveglianza, non solo in ambito domestico, ma anche in alberghi, centri commerciali, stazioni ferroviarie, aeroporti, ospedali, case di riposo e luoghi di lavoro.

30. Tra i principali “assistenti personali” (*personal* - o *digital* o *smart* o *virtual* - *assistants*) installabili su dispositivi mobili, attualmente si possono citare: ClickUp, Socratic, Amazon Alexa, Cortana, Wally, Tabnine, Parrot IA, S Voice, Siri, Google Assistant. Per una sintetica presentazione delle funzioni e delle potenzialità di questi strumenti, ma anche dei rischi connessi al loro uso, si veda, oltre a *Assistente virtuale*, Voce in “Wikipedia. L’enciclopedia libera”, MIALICH 2021. In merito al problema dei rapporti tra assistenti digitali e tutela dei dati personali si veda GARANTE PER LA PROTEZIONE DEI DATI PERSONALI 2021.

concretamente disponibili numerosi strumenti atti a gestirli, accorparli e organizzarli.

2.3. La “Teoria della mente” e la nuova configurazione del rapporto uomo-computer

In base alle osservazioni svolte è chiaro che l’interazione col computer, come risultato del flusso d’informazioni che a esso vengono trasferite – volontariamente o involontariamente, consapevolmente o inconsapevolmente –, induce comunque la costruzione d’un vero e proprio “specchio digitale” dell’utente, considerato come mera entità di provenienza. A sua volta, poi, questo modello che viene generato, rimanendo continuamente aggiornabile, modificabile e incrementabile, diviene in qualche modo la base dei futuri “comportamenti” del computer nei confronti dell’utente stesso. Come nelle relazioni interpersonali, azioni e reazioni sono governate dai “modelli mentali” autonomamente e reciprocamente sviluppati dalle parti, allo stesso modo l’utente deve prendere coscienza del fatto che, nell’interazione col computer, gradatamente egli costruisce, in esso, un’immagine di sé; sulla base di questa, poi, gli algoritmi d’apprendimento di cui il dispositivo elettronico è dotato svilupperanno “comportamenti” ritenuti adeguati, in quanto, appunto, corrispondenti al modello che al suo interno si è venuto configurando.

A una comprensione più approfondita di questo tema – d’importanza centrale sotto molti e diversi profili, da quello economico e commerciale (basti pensare all’elaborazione delle strategie produttive, all’impostazione delle campagne pubblicitarie e, quindi, all’orientamento dei consumi) fino a quello psicologico, sociologico e giuridico (sono particolarmente rilevanti in questo ambito i fenomeni connessi allo sviluppo esponenziale del social network, ma anche le grandi trasformazioni in atto nelle scienze sociali e, in particolare, nel diritto) – giova il riferimento a quella che gli psicologi chiamano “Teoria della mente”³¹.

L’espressione “Teoria della mente” (ToMM – *Theory of Mind Mechanisms*) è utilizzata sovente con significati diversi: a livello astratto, nell’ambito della “filosofia della mente”, essa individua un

modello ontologico e strutturale dei processi mentali, facendo esplicitamente riferimento agli studi sulla “natura della mente”. All’interno della “psicologia cognitiva”, invece, con questa stessa espressione si designano le analisi volte a configurare modelli del “funzionamento della mente”.

Evitando di addentrarsi in indagini complesse ed estese, in cui sono coinvolte numerose discipline tra loro distinte e a volte lontane (non solo la filosofia, la psicologia, la psicoterapia, la pedagogia e la psicanalisi, ma anche la neurologia e le scienze cognitive in generale), in questa sede può ritenersi sufficiente assumere il significato e la funzione della “Teoria della mente” restringendoli e focalizzandoli su due concetti fondamentali: la “mentalizzazione” e la “metacognizione”. La “mentalizzazione” è la facoltà che ci consente di rappresentare (e rappresentarci), di “tenere a mente” gli stati d’animo, i pensieri, riferibili agli altri e a noi stessi. Più ampia è, invece, l’estensione semantica del termine “metacognizione”, che designa la capacità – molto probabilmente esclusiva degli esseri umani – di cercare di comprendersi in termini di stati mentali (quindi, non solo pensieri, ma anche sentimenti, desideri, credenze), in modo da poter attribuire significato all’esperienza e, quindi, riuscire in qualche modo a configurare in anticipo le reciproche azioni, sulla base della previsione dei comportamenti.

Nell’analisi dei rapporti uomo-computer, prendendo atto della rivoluzione tecnologica in corso in questi ultimi anni, soprattutto nel campo dell’Intelligenza artificiale e dell’apprendimento automatico, a questo punto si aprono scenari assolutamente nuovi e in gran parte imprevedibili: il computer non è più la macchina inerte che si limita a eseguire, in modo rigido e sequenziale, gli ordini impartitigli dal programmatore; essendo dotato di “capacità di apprendimento”, esso diviene interlocutore attivo e, per certi versi, imprevedibile, in grado non solo di registrare enormi masse di dati, ma anche di elaborarle in modo da estrarne nuova conoscenza (*data mining*) e, quindi, di acquisire esperienza, sulla cui base formulare proposte d’azione o prendere decisioni.

I suggerimenti che il computer fornisce all’utente in ordine all’acquisto di beni e servizi online

31. Sulla “Teoria della mente” esiste ormai una imponente letteratura. Oltre a *Teoria della mente*, Voce in “Wikipedia. L’enciclopedia libera”, tra le opere più recenti e complete ci si limita qui a segnalare: CAMAIONI 2016; LECCE–CAVALLINI–PAGNIN 2010.

(il libro, il film, il brano musicale, l'albergo o l'agenzia immobiliare da selezionare) non sono esplicitamente predisposti per lo specifico utente che volta a volta effettua la consultazione, né sono stati elaborati in anticipo, in relazione a categorizzazioni teoriche più o meno ampie e comprensive. L'assoluta novità introdotta dalla tecnologia del *machine learning* è rappresentata proprio dalla caratteristica intrinseca degli algoritmi di apprendimento, in quanto capaci di sostituire alla necessità dell'analisi minuziosa di quantità indefinite di dati la possibilità di estrarre conoscenza ed elaborare esperienza sulla base di un numero ristretto di precedenti casi concreti³².

3. Il problema della regolazione giuridica dell'Intelligenza artificiale: l'approccio europeo

L'approccio europeo all'Intelligenza artificiale fin dall'inizio è stato concepito partendo dalla considerazione dell'importanza centrale dell'IA nello sviluppo complessivo della società e nella configurazione delle modalità e dello stile di vita dell'uomo³³. Si è efficacemente affermato che il modo in cui si affronteranno l'introduzione, l'applicazione e la diffusione di queste nuove tecnologie non semplicemente condizionerà, ma addirittura "disegnerà" il mondo del futuro. La strategia europea per l'Intelligenza artificiale si è posta, dunque, come obiettivo la trasformazione dell'Europa in un centro di riferimento a livello mondiale, non

solo per la progettazione e lo sviluppo di soluzioni innovative, ma anche per la loro progressiva sperimentazione e adozione, tanto nel settore pubblico quanto in quello privato.

In ordine al conseguimento di questo risultato, già nella sua Comunicazione sul tema "Intelligenza artificiale per l'Europa"³⁴ la Commissione europea aveva evidenziato la necessità di: a) aumentare la capacità tecnologica e industriale dei diversi Stati dell'Unione incrementando la ricerca e l'innovazione, oltre che migliorando l'accesso ai dati; b) prepararsi ai cambiamenti socio-economici che ne deriveranno, incoraggiando la modernizzazione dei sistemi d'istruzione e formazione; c) anticipare e sostenere le transizioni e i cambiamenti nel mercato del lavoro adattando i sistemi di protezione sociale; d) infine, garantire un quadro etico e giuridico adeguato, basato sui principi e i valori propri dell'Unione e, quindi, in linea con la sua Carta dei diritti fondamentali³⁵.

Di fronte al rapido sviluppo tecnologico cui si sta assistendo negli anni più recenti, all'interno di un contesto politico internazionale in cui un numero sempre crescente di Stati investe massicciamente nell'Intelligenza artificiale, l'Unione europea si è trovata nella necessità di affrontare con urgenza il problema della sua credibilità e competitività, avendo riguardo a un ampio orizzonte temporale, in relazione, da una parte, alla molteplicità e rilevanza delle iniziative prese in questo campo soprattutto dalle grandi potenze mondiali (in particolare: Stati Uniti, Cina, Giappone e Canada), dall'altra, al

32. In questa direzione l'obiettivo che ci si pone è quello d'imparare da un solo esempio o da pochi esempi. OpenAI, l'organizzazione non-profit fondata da Elon Musk, ha diffuso un sistema di IA che consente a un robot di apprendere come svolgere un determinato compito semplicemente osservando un essere umano intento a farlo per una sola volta. Com'è facile intuire, si tratta di un progresso notevole, per quanto ancora siano necessari ulteriori affinamenti per arrivare a costruire macchine in grado di eseguire incarichi sempre più complessi senza la necessità di lunghi periodi di addestramento. L'apprendimento per imitazione è stato solitamente applicato per risolvere differenti compiti presi singolarmente. Di solito questo richiede un attento lavoro ingegneristico o un significativo numero di esempi; idealmente, invece, i robot dovrebbero essere capaci d'imparare da poche dimostrazioni, applicando poi il principio generale a nuove situazioni dello stesso tipo, senza che si richieda un lavoro ulteriore.

33. Si veda Commissione europea, *Approccio europeo all'Intelligenza artificiale*.

34. Si veda Comunicazione della Commissione, COM (2018) 237, *Intelligenza artificiale per l'Europa*, 25 aprile 2018.

35. La Carta dei diritti fondamentali dell'Unione Europea (216/C 202/02), conosciuta in Italia anche come "Carta di Nizza", è stata pubblicata nella Gazzetta ufficiale dell'Unione europea il 7 giugno 2016. Proclamata ufficialmente a Nizza nel dicembre del 2000 dal Parlamento europeo, dal Consiglio dell'Unione europea e dalla Commissione, è diventata poi giuridicamente vincolante con l'entrata in vigore del Trattato di Lisbona, nel dicembre del 2009; attualmente essa ha lo stesso effetto giuridico dei Trattati dell'Unione.

continuo e rapido sviluppo della ricerca pubblica e privata in un settore d'importanza strategica centrale. È divenuta, quindi, una priorità assoluta per l'Ue preservare e incrementare le proprie risorse in modo da poter creare un ambiente atto a stimolare gli investimenti, utilizzando anche i finanziamenti pubblici per incentivare gli investimenti privati. A questo primo obiettivo si affianca, però, l'esigenza di analizzare e affrontare i molteplici rischi che dalla sempre più vasta diffusione di queste tecnologie derivano per la tutela della salute e della sicurezza, oltre che per la garanzia dei diritti fondamentali dei cittadini.

In entrambe queste due direzioni, nel 2021 la Commissione europea si è mossa formulando una "Proposta per un quadro normativo sull'Intelligenza artificiale"³⁶ e un "Piano coordinato sulla revisione dell'Intelligenza artificiale"³⁷, essenzialmente basato su una forte collaborazione tra gli Stati membri, invitati a sviluppare e aggiornare le loro strategie nazionali in ordine alla realizzazione di una leadership globale dell'Europa in questo importante settore. Nel corso del mese di giugno 2023 il Parlamento europeo ha fissato la propria posizione negoziale in merito alla normativa sull'IA (AI Act), come primo tentativo al mondo di elaborare e strutturare un insieme di regole volte a disciplinare lo sviluppo e l'applicazione di queste nuove tecnologie. Da ultimo, poi, nel gennaio del 2024 la Commissione ha pubblicato una "Comunicazione in materia di promozione delle start-up e dell'innovazione in una IA affidabile"³⁸ in cui, rifacendosi ai Piani d'azione coordinati del 2018 e del 2021, definisce un quadro strategico d'investimenti per un'Intelligenza artificiale affidabile affinché l'Unione possa capitalizzare le sue risorse, in particolare la sua infrastruttura di supercalcolo, leader

a livello mondiale, e promuovere un ecosistema europeo innovativo in cui le start-up e gli innovatori possano collaborare a stretto contatto con gli utilizzatori industriali e attrarre investimenti.

In merito è importante rilevare che, nella citata Comunicazione, la Commissione mostra un sostanziale cambiamento di politica nei confronti dell'IA generativa, adeguandosi ai più recenti sviluppi della tecnologia. Di conseguenza, la Legge europea sull'Intelligenza artificiale (AI Act)³⁹, entrata in vigore il 1° agosto 2024, include anche disposizioni sull'IA generativa, ampliando la Proposta originaria presentata dalla Commissione nell'aprile del 2021.

Lo sviluppo e l'impiego sempre più esteso dell'IA possono rappresentare un elemento di mutazione radicale nella qualità della vita, a livello sia individuale che sociale, in positivo, ma anche in negativo. È dunque evidente che la prosperità e la crescita economica dell'Europa dipenderanno strettamente dal modo in cui saranno utilizzati i dati e le tecnologie della connessione. Finora, però, secondo quanto evidenziato in una recente nota della Corte dei conti europea⁴⁰, l'Ue ha mostrato di conseguire scarsi risultati nello sviluppare un ecosistema di Intelligenza artificiale europeo; in particolare, non è riuscita a far accelerare gli investimenti nel settore a un ritmo paragonabile a quello dei *competitors* mondiali. Fin dal 2018 la Commissione europea si è impegnata in numerose iniziative, affrontando anche questioni di fondo, quali la regolazione, le infrastrutture, la ricerca e gli investimenti. Tuttavia, le misure prese non sono state adeguatamente coordinate con quelle degli Stati membri; in particolare, non è stato condotto in maniera sistematica il monitoraggio degli investimenti. Per il futuro è, dunque, da considerare imprescindibile il ricorso a

36. Vedi Commissione europea, Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza artificiale (Legge sull'Intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, COM/2021/206, 21 aprile 2021.

37. Vedi COMMISSIONE EUROPEA 2021. Il Piano coordinato, inizialmente pubblicato nel 2018, è stato poi aggiornato nel 2021, mentre è del 2024 la comunicazione della Commissione in materia di promozione delle start-up e innovazione nell'IA affidabile, di cui alla nota successiva.

38. DOC. COM(2024)28, 24 gennaio 2024.

39. L'AI Act stabilisce regole specifiche per i modelli di AI per finalità generali (Capo V), ponendo obblighi più stringenti per le ipotesi in cui si tratti di modelli classificati come "con rischio sistemico" (art. 5 e ss.); tali norme entreranno in vigore nell'agosto del 2025.

40. Si veda la nota rilasciata dalla CORTE DEI CONTI EUROPEA 2024. Cfr. anche PEDRAZZOLI 2024.

una governance più forte e a investimenti pubblici e privati più mirati e consistenti.

A livello globale, nella corsa agli investimenti nell'IA, l'Unione sta incontrando numerosi ostacoli. Dal 2015 gli investimenti effettuati in strumenti di capitale di rischio sono stati molto inferiori a quelli degli Stati Uniti e della Cina; si stima che il divario tra gli Stati Uniti e l'Ue in termini di investimenti complessivi nell'IA sia più che raddoppiato tra il 2018 e il 2020. In tale contesto l'Ue ha progressivamente adottato misure atte a sviluppare al proprio interno un quadro di coordinamento orientato verso l'incremento degli investimenti e l'elaborazione di un'adeguata normativa. Come si è osservato, però, i piani predisposti dalla Commissione per l'IA nel 2018 e nel 2021, pur essendo esaustivi e ampiamente in linea con le migliori pratiche internazionali, hanno registrato forti ritardi nell'attuazione, mentre gli obiettivi d'investimento dell'Ue risultano ormai troppo vaghi e obsoleti, non essendo stati più aggiornati dal 2018. La Commissione è riuscita ad aumentare la spesa a carico del bilancio dell'Ue per progetti di ricerca nel settore dell'IA, ma non ha potenziato significativamente il cofinanziamento privato e non si è impegnata a sufficienza per garantire che i progetti di ricerca finanziati siano poi adeguatamente commercializzati e sfruttati.

3.1. La diffusione dell'Intelligenza artificiale: opportunità e rischi

Gli scenari che si aprono in conseguenza dell'introduzione sempre più estesa dell'IA nelle strutture sociali come nella vita privata dei cittadini sono estremamente vari e complessi. Se da un lato, infatti, risulta fin d'ora evidente che l'Intelligenza artificiale può rendere disponibili veicoli e mezzi di trasporto più sicuri, contribuire a migliorare la sanità, l'istruzione e la formazione, aumentare la

sicurezza nello svolgimento di attività lavorative pericolose, rendere più agevole l'accesso all'informazione e, in generale, consentire lo sviluppo di nuovi servizi per i cittadini, sotto un diverso profilo è però chiaro che si sono aperte prospettive ampiamente inesplorate anche per quanto attiene al mondo delle imprese e della produzione: sarà infatti possibile sviluppare una nuova generazione di prodotti e servizi anche in quei settori in cui le aziende europee sono già presenti in una posizione di forza (come l'economia circolare⁴¹, l'agricoltura, la sanità, la moda, il turismo), ma anche rendere più sicura la manutenzione dei macchinari, ottenere risparmi di energia, aumentare sia la produzione che la qualità, ottenere percorsi di vendita più fluidi ed efficienti, infine, offrire servizi migliori e su misura per i clienti.

Altri benefici dell'IA riguarderanno, poi, il settore dei servizi pubblici, in cui si renderanno possibili nuove opzioni, in particolare per la gestione dell'energia e lo smaltimento dei rifiuti, in tal modo contribuendo anche al conseguimento degli obiettivi del "Green Deal"⁴² europeo. Inoltre, con l'impiego di queste nuove tecnologie la democrazia risulterà rafforzata dalla possibilità di effettuare verifiche basate sui dati, combattere la disinformazione⁴³, prevenire gli attacchi informatici e, in ultima analisi, garantire l'accesso a informazioni di qualità. Sul piano della sicurezza, in particolare, l'IA costituirà un ausilio prezioso nella prevenzione dei reati. Come supporto nell'esercizio della giustizia penale, poi, essa consentirà di elaborare molto velocemente grandi quantità di dati, in modo da poter prevedere e prevenire crimini e attacchi terroristici, ma anche individuare e contrastare pratiche illegali o inappropriate della rete; in caso di crimini informatici potrà essere utilizzata sia per la difesa, sia per predisporre strategie di attacco contro obiettivi chiave⁴⁴.

41. Si veda in PARLAMENTO EUROPEO 2023: "L'economia circolare è un modello di produzione e consumo che implica condivisione, prestito, riutilizzo e riparazione, ricondizionamento e riciclo dei materiali e prodotti esistenti il più a lungo possibile. In questo modo si estende il ciclo di vita dei prodotti, contribuendo a ridurre i rifiuti al minimo"; cfr. anche Commissione europea 2020.

42. Per una sintetica ma efficace presentazione, si veda Commissione europea, *Il Green Deal europeo. Per diventare il primo continente a impatto climatico zero*, 2020.

43. Su questo aspetto particolare si vedano le pagine del sito web del Parlamento europeo dedicate al tema *Disinformazione*.

44. Si veda CONSIGLIO EUROPEO-CONSIGLIO DELL'UNIONE EUROPEA 2024. Il Centro europeo per la criminalità informatica (European Cybercrime Centre - EC3), istituito nel 2013, ha fornito un importante contributo alla

Ai sempre più rilevanti e diffusi benefici connessi all'impiego delle tecnologie proprie dell'Intelligenza artificiale corrispondono però anche numerosi rischi, attinenti in particolare: al loro possibile abuso o sottoutilizzo, alla disciplina della responsabilità civile per danni causati da dispositivi o servizi che incorporino funzioni specifiche dell'IA, ma anche alle esigenze di tutela dei diritti fondamentali dell'Unione europea. Gravi minacce, poi, possono sorgere nei confronti della protezione dei dati e per il diritto alla vita privata, come pure per l'esercizio della democrazia e la concorrenza del mercato. Ancora, può essere messa seriamente a repentaglio l'integrità fisica della persona qualora si tratti di applicazioni di IA collocate a contatto o addirittura integrate nel corpo umano, mentre un uso non adeguatamente regolamentato di queste nuove tecnologie negli armamenti potrebbe condurre alla perdita del controllo su di essi, con inimmaginabili conseguenze sugli equilibri internazionali e sulla pacifica coesistenza tra gli Stati.

Molteplici e gravi possono essere, infine, i problemi derivanti dalla mancanza di trasparenza nella struttura e nel funzionamento dei sistemi di IA: le informazioni acquisite possono essere insufficienti, errate o addirittura volutamente distorte, inducendo l'utente a compiere scelte fuorvianti o a prendere decisioni del tutto sbagliate. A seguito dei più recenti sviluppi tecnologici, inoltre, può accadere anche che l'utente non sia in grado di comprendere se stia interagendo con una persona umana ovvero con un sofisticato sistema conversazionale intelligente.

3.2. Il contesto geopolitico e gli obiettivi della politica europea per l'IA

La complessità delle valutazioni relative al necessario bilanciamento tra opportunità e svantaggi nell'introduzione massiccia dell'IA deve poi

integrarsi all'interno d'un contesto geopolitico fortemente condizionato dalla presenza dominante dei due principali *competitors* globali: gli Stati Uniti e la Cina. Entrambe orientate verso una decisa incentivazione degli investimenti pubblici e privati nell'IA, queste due potenze hanno in comune il riconoscimento dell'importanza strategica fondamentale del settore, ma anche la preoccupazione di non ostacolarne in alcun modo le dinamiche e lo sviluppo con un apparato normativo di tipo analitico e cautelativo, caratterizzandosi, i primi, per i timori relativi alla tutela dei diritti fondamentali e, la seconda, per l'impostazione disinvoltamente autoritaria di una politica irrispettosa tanto della privacy dei cittadini quanto delle esigenze di trasparenza dei sistemi⁴⁵.

Nel quadro normativo statunitense, in materia di regolazione dell'IA è considerata centrale l'importanza della sicurezza e dell'affidabilità dell'innovazione tecnologica in corso. È del 30 ottobre 2023 l'*Executive Order* del Presidente Biden⁴⁶, in cui sono enunciati gli otto principi ispiratori della politica che s'intende attuare. A livello federale esistono poi iniziative di legislazione c.d. "orizzontale", relativa cioè a una molteplicità di sistemi di IA rilevanti sul piano decisionale e con possibili ricadute sulla stessa vita degli utenti, ma anche atti di regolazione dell'Intelligenza artificiale generativa, della tecnologia delle risorse umane, nonché delle piattaforme e delle comunicazioni online⁴⁷.

Sempre al di fuori dell'Unione europea è qui da segnalare anche la posizione del Regno Unito che, nel 2022, è stato il terzo paese al mondo, dopo gli Stati Uniti e la Cina, a registrare i maggiori investimenti privati e il più alto numero di nuove imprese attive nel settore dell'IA, distinguendosi per l'importanza delle ricerche in svolgimento e l'entità delle risorse pubbliche destinate (*Global AI Index*, Tortoise Media, 2022). L'approccio espresso è manifestamente orientato verso una regolazione

lotta contro la criminalità informatica concorrendo a proteggere i cittadini, le imprese e i governi europei dalla criminalità online.

45. Si vedano, fra gli altri: BURT 2021; AARONSON 2020. Per quanto riguarda la Cina occorre fare riferimento al documento pubblicato dal Consiglio di Stato cinese il 20 luglio 2017, che definisce una progettazione d'alto livello e illustra l'approccio del paese allo sviluppo della tecnologia e delle applicazioni dell'IA, stabilendo obiettivi generali validi fino al 2030. Cfr. CONSIGLIO DI STATO CINESE 2017. Si vedano, fra gli altri: ROBERTS-COWLS-MORLEY et al. 2021; KHANAL-ZHANG-TAEIHAGH 2024.

46. Si veda SULMICELLI 2023.

47. Si veda ANAND 2024.

prudente ma flessibile, attenta all'esigenza di non limitare o imbrigliare gli sviluppi della nuova tecnologia, fornendo indicazioni chiare e tempestive all'industria e nel contempo promuovendo l'interoperabilità a livello internazionale. In questa prospettiva la regolazione, pur essendo considerata d'importanza primaria, non si configurerebbe sempre come lo strumento più efficace per promuovere e sostenere un'innovazione responsabile, essendo invece necessario il ricorso a varie forme d'integrazione, basate sull'osservanza di principi fondamentali come l'equità, la sicurezza e la protezione, la trasparenza e la contestabilità, la responsabilità e la governance. All'implementazione di tali principi dovranno provvedere le Autorità già esistenti, mentre eventuali modifiche e aggiustamenti potranno derivare dall'esecuzione di test e sperimentazioni in ambienti controllati, ma anche dallo svolgimento di consultazioni pubbliche.

Il 1° aprile 2024, gli Stati Uniti e il Regno Unito hanno firmato un *memorandum of understanding* sull'Intelligenza artificiale⁴⁸, che costituisce il primo accordo bilaterale al mondo relativo alla valutazione e allo studio dei rischi connessi all'IA. In forza dell'accordo raggiunto Stati Uniti e Regno Unito collaboreranno mettendo reciprocamente a disposizione le proprie conoscenze scientifiche in ordine all'implementazione di modelli di IA sicuri: in concreto l'US AI Safety Institute (US AISI) e l'omologo Istituto britannico potranno condividere informazioni attinenti alla sicurezza dei modelli di IA e quindi ai rischi a essi associati, in modo da sviluppare alcuni approcci metodologici volti ad affrontarli e limitarli.

L'analisi, sia pur sommaria, del contesto geopolitico in cui l'Unione si trova inserita consente di comprendere agevolmente per quali motivi l'Europa abbia dovuto affrontare con urgenza il problema di realizzare una disciplina normativa del fenomeno complessivo dell'IA: si tratta di entrare in una difficile, ma necessaria competizione con i Paesi tecnologicamente più avanzati, creando un

ambiente favorevole alla ricerca e allo sviluppo, assicurando e stimolando i necessari investimenti, ma nel contempo mantenendo salda l'esigenza di riaffermare e tutelare quei diritti fondamentali, quei valori e quei principi su cui l'Europa stessa si fonda.

L'obiettivo principale della politica europea è dunque duplice: da una parte, garantire all'UE la possibilità di acquisire e mantenere una leadership mondiale in questo campo, facendo però in modo che l'impiego dell'IA avvenga nel pieno rispetto dei diritti fondamentali, dei principi e dei valori europei; dall'altra, assicurarsi che la regolazione europea dell'IA risulti equilibrata, idonea a temperare concezioni ed esigenze diverse (talora perfino tra loro contrastanti), proporzionata rispetto ai rischi da evitare o limitare, atta a incoraggiare e promuovere lo sviluppo tecnologico suscitando e garantendo la fiducia di cittadini e utenti, flessibile e adattabile nei confronti dei rapidi e continui cambiamenti che caratterizzano queste tecnologie, assicurando nel contempo anche quel grado di certezza e prevedibilità che è essenziale in un settore tanto strategico e delicato. Si tratta, evidentemente, di un equilibrio non facile da trovare a livello nazionale né, tantomeno, a livello europeo (o globale): le regole sono necessarie per assicurare il rispetto dei diritti e dei valori su cui si basa l'Unione europea, ma non devono costituire un ostacolo rispetto a quello sviluppo scientifico, economico e sociale che l'IA può produrre⁴⁹.

Agli obiettivi fin qui chiariti deve ovviamente corrispondere un ambito d'applicazione particolarmente esteso: per esplicare concretamente la loro efficacia le nuove norme dovranno applicarsi non solo all'immissione sul mercato, ma anche alla messa in servizio e all'uso dei "sistemi di IA", intesi, in senso ampio, come software sviluppati con tecniche specificamente indicate e idonei a generare contenuti, previsioni, raccomandazioni o decisioni, tali da poter influenzare l'ambiente con cui interagiscono⁵⁰. Le nuove norme dovranno poi essere

48. Si veda U.S. DEPARTMENT OF COMMERCE, OFFICE OF PUBLIC AFFAIRS 2024.

49. Per un confronto tra la strategia USA e quella europea cfr. CHITI–MARCHETTI 2020; CATH–WATCHER–MITTELSTADT et al. 2018.

50. Nel testo inglese della Proposta di Regolamento la definizione dei "sistemi di IA" è così espressa: "Artificial Intelligence system (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with".

applicate direttamente, allo stesso modo, in tutti gli Stati membri, sulla base di una definizione dell'Intelligenza artificiale⁵¹ sufficientemente flessibile e comprensiva, atta ad esprimerne la complessità concettuale e le potenzialità di sviluppo, ma allo stesso tempo accompagnandone la rapida evoluzione.

4. La Legge europea sull'Intelligenza artificiale (AI Act): entrata in vigore e ambito di applicazione

Il 13 marzo 2024 il Parlamento europeo ha approvato, con ampia maggioranza, il Regolamento

sull'Intelligenza artificiale (AI Act)⁵²; a poco più di due mesi di distanza, il 21 maggio successivo, è giunta anche l'approvazione del Consiglio dell'Unione europea⁵³. La pubblicazione nella Gazzetta Ufficiale dell'Unione è avvenuta il 12 luglio 2024⁵⁴.

Dopo venti giorni, in data 1° agosto 2024, la legge è entrata finalmente in vigore, ma comincerà a essere applicata solo dopo 24 mesi⁵⁵. Sono comunque previste alcune eccezioni: le regole relative ai sistemi di IA vietati si applicano 6 mesi dopo l'entrata in vigore; i codici di buone pratiche, 9 mesi dopo; le norme sui sistemi di IA per finalità

51. La definizione dell'Intelligenza artificiale, elaborata dal "Gruppo indipendente di esperti di alto livello sull'Intelligenza artificiale" (*High Level Expert Group on Artificial Intelligence* – AI HLEG) nominato dalla Commissione europea, era così espressa: "I sistemi di intelligenza artificiale (IA) sono sistemi software (ed eventualmente hardware) progettati dall'uomo che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il proprio ambiente attraverso l'acquisizione di dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulle conoscenze, o elaborando le informazioni derivate da questi dati e decidendo le migliori azioni da intraprendere per raggiungere l'obiettivo dato. I sistemi di IA possono usare regole simboliche o apprendere un modello numerico, e possono anche adattare il loro comportamento analizzando come l'ambiente è influenzato dalle loro azioni precedenti".

52. Il testo del Regolamento contenente le proposte di emendamento, frutto dell'accordo raggiunto con gli Stati membri nel dicembre 2023, dopo avere ottenuto l'approvazione del "Comitato dei rappresentanti permanenti degli Stati membri" (COREPER), della "Commissione per le libertà civili, la giustizia e gli affari interni" (LIBE) e della "Commissione per il mercato interno e la protezione dei consumatori" (IMCO), è stato approvato dai deputati europei con 523 voti favorevoli, 46 contrari e 49 astensioni.

53. Il testo del "Regolamento recante norme armonizzate sull'Intelligenza artificiale" (Legge sull'Intelligenza artificiale) è stato approvato, insieme, dal Parlamento europeo e dal Consiglio dell'Unione europea in data 21 maggio 2024.

54. Il testo pubblicato nella Gazzetta Ufficiale dell'Unione europea il 12 luglio 2024 reca il titolo "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)". La cronologia storica dell'AI Act è presentata nella [Historic Timeline](#) del Future of Life Institute. Per una presentazione complessiva della legge si vedano, tra i tanti: ALAIMO 2023; BARBAJA 2024; BOCCHINI 2024; CAPACCI-GALLI-LOREGGIA-MAROCCIA 2024; IASELLI 2024; LALLI 2024, VOIGT-HULLEN 2024. Per agevolare l'analisi del testo – molto esteso, comprendendo 180 considerando, 113 articoli distribuiti in 13 Capi e 13 Allegati – sono stati già implementati alcuni utili strumenti di navigazione. Tra questi si possono qui citare: l'*AI Act Explorer*, sviluppato dal "Future of Life Institute"; [ClaiRK](#), strumento articolato che consente d'effettuare anche ricerche testuali e semantiche, realizzato dal "Digital Policy Alert team"; Lexroom AI Act, nuovo modulo della piattaforma Lexroom.ai, progettato da una giovane start-up innovativa in modo da poter rispondere a domande in linguaggio naturale in merito ai contenuti della legge europea sull'Intelligenza artificiale, fornendo in pochissimo tempo una "bozza di parere" o una "risposta pragmatica", ma anche citando e allegando le fonti giuridiche utilizzate. Infine, un indice interattivo dell'AI Act è stato prodotto anche dal MIAI (*Multidisciplinary Institute in Artificial Intelligence*) dell'Université Grenoble Alpes.

55. Il Regolamento europeo sull'IA, come tutti i regolamenti dell'Ue, è un atto normativo di carattere generale, vincolante e direttamente applicabile; essendo *self executive*, non richiede alcuna legge di recepimento, ma entra direttamente a far parte del corpo legislativo nazionale di ciascuno Stato membro.

generali (*General Purpose AI* – GPAI), compresa la governance, 12 mesi dopo; infine, gli obblighi per i sistemi di IA ad alto rischio entreranno in vigore solo dopo tre anni, mentre il periodo di tolleranza è esteso a ben quattro anni per i sistemi IA ad alto rischio che siano in uso da parte di pubbliche autorità e immessi sul mercato prima dell'entrata in vigore della legge. Gli Stati membri, poi, hanno 12 mesi di tempo per creare un'Autorità locale di notifica e di sorveglianza del mercato, ma potranno disporre di due anni per approntare le regolamentazioni locali⁵⁶.

Come già precisato, l'AI Act tende a garantire che i sistemi di IA immessi sul mercato europeo siano sicuri e rispettino i diritti e i valori fondamentali dell'Unione europea; nel contempo, però, esso si propone anche di stimolare gli investimenti e l'innovazione nel settore dell'Intelligenza artificiale, agevolando lo sviluppo di un mercato unico per applicazioni di IA lecite, sicure e affidabili. Norme nazionali diverse in ognuno dei singoli Stati membri, infatti, porterebbero alla frammentazione del mercato interno, mettendone a rischio il vantaggio competitivo e pregiudicando la necessaria certezza del diritto nei confronti di quanti sviluppano, producano, utilizzino o importino sistemi di IA all'interno dell'Unione.

A livello comunitario, relativamente a una materia tanto complessa oltre che in continua evoluzione, si rende evidente la ferma volontà d'inserire nella legislazione europea una disciplina organica e unitaria, tale da poter diventare non solo un punto di riferimento cogente per le legislazioni degli altri Stati membri, ma anche un

prototipo normativo suscettibile d'essere assunto come modello cui ispirarsi per i Paesi al di fuori dell'Unione.

Il testo del Regolamento approvato dal Parlamento europeo e dal Consiglio dell'Unione conferma l'approccio basato sul rischio, già adottato nelle precedenti versioni. Pertanto, sono imposti obblighi diversificati in base al livello di rischio che ciascuna delle categorie in cui si possono classificare i sistemi di IA può comportare nei confronti dei diritti e delle libertà personali: da un rischio inaccettabile si passa alla configurazione di un rischio alto e, successivamente, basso o minimo.

Rispetto alla Proposta presentata dalla Commissione il 21 aprile 2021⁵⁷, nel quadro normativo che emerge dal testo definitivo della legge sull'Intelligenza artificiale sono principalmente evidenziabili: le nuove disposizioni relative ai modelli di IA a carattere generale ad alto impatto; la ridistribuzione dei poteri in ambito istituzionale, con una più forte centralizzazione delle decisioni esecutive a livello sovranazionale; l'introduzione di nuove pratiche vietate connesse all'uso di sistemi di IA; le precisazioni introdotte in riferimento alle situazioni in cui l'impiego di sistemi d'identificazione biometrica remota in tempo reale in spazi accessibili al pubblico si rende strettamente necessario per il contrasto alla criminalità; infine, la valutazione d'impatto, preliminare all'attivazione di un sistema di IA.

Le regole armonizzate stabilite nel Regolamento sono state concepite in modo da non pregiudicare la normativa vigente dell'Unione, in particolare in materia di diritti fondamentali, protezione dei

56. Si riporta qui di seguito lo scadenario dettagliato previsto dal legislatore per l'entrata in vigore e l'applicazione della legge (art. 113): l'entrata in vigore della legge sull'AI è avvenuta 20 giorni dopo la pubblicazione nella Gazzetta ufficiale dell'Unione europea, il 1° agosto 2024; dal 2 febbraio 2025 si applicheranno i Capi I (disposizioni generali) e II (sistemi di IA vietati); dal 2 agosto 2025 si applicheranno il Capo III, Sezione 4 (Autorità di notifica), il Capo V (modelli di IA per scopi generali), il Capo VII (governance), il Capo XII (riservatezza e sanzioni) e l'articolo 78 (riservatezza), ad eccezione dell'articolo 101 (sanzioni per i fornitori di IA *general purpose*); dal 2 agosto 2026 si applicherà l'intera legge sull'IA, ad eccezione dell'articolo 6 (regole di classificazione per i sistemi di IA ad alto rischio), paragrafo 1, e degli obblighi corrispondenti; dal 2 agosto 2027 si applicheranno l'articolo 6, paragrafo 1, e gli obblighi corrispondenti. Per un'analisi dettagliata si veda ZORLONI 2024.

57. La Proposta di Regolamento sull'Intelligenza artificiale presentata dalla Commissione europea il 21 aprile 2021 [COM(2021) 206], mirando a stabilire regole comuni atte a salvaguardare i diritti umani fondamentali e allo stesso tempo agevolare lo sviluppo del mercato interno, forniva una definizione e classificazione dei sistemi di IA, indicava i requisiti obbligatori per la loro progettazione e realizzazione, commercializzazione e utilizzo, prevedendo anche diversi strumenti di controllo e uno specifico apparato sanzionatorio.

dati, tutela dei consumatori, sicurezza dei prodotti, occupazione e protezione dei lavoratori.

Sotto il profilo soggettivo, in base all'art. 2 la legge si applica agli utilizzatori dei sistemi di IA, siano essi enti pubblici o privati, ma anche ai fornitori, ai *deployer*, agli importatori, ai distributori, ai fabbricanti, ai rappresentanti autorizzati e, in generale, a quanti siano comunque coinvolti nell'uso di tali sistemi. Nel successivo art. 3 sono fornite le definizioni relative alle figure di "fornitore" ("provider"), "deployer", "rappresentante autorizzato", "importatore", "distributore" e "operatore". Sul piano pratico è qui da rilevare l'importanza della corretta individuazione della qualifica attribuibile al soggetto, in quanto da essa discende l'attribuzione d'un insieme di obblighi e responsabilità. In particolare, com'è stato osservato⁵⁸, in considerazione del fatto che tali obblighi e responsabilità sono per la maggior parte normativamente previsti a carico del provider, la distinzione tra la figura di "provider" e quella di "deployer" – che sovente può risultare difficile a causa del notevole livello di personalizzazioni che le aziende tendono ad apportare ai sistemi – e la conseguente corretta attribuzione della corrispondente qualifica comportano conseguenze pratiche rilevanti.

Per quanto attiene al criterio della territorialità, l'AI Act ha adottato una soluzione simile a quella utilizzata dal GDPR e da altre normative in materia di tecnologie digitali: esso, infatti, trova applicazione non solo con riferimento ai fornitori di sistemi di IA che li immettono sul mercato o li utilizzano nel territorio dell'Unione europea, ma anche nei confronti di quanti operano al di fuori di tale ambito se i risultati generati da tali sistemi sono poi comunque utilizzati in Europa. La *ratio* della norma, dunque, estendendo la sua applicabilità anche a soggetti esterni all'Unione, funziona come misura protettiva, nel senso di prevenire l'introduzione e la commercializzazione nel mercato dell'Ue di sistemi di Intelligenza artificiale che abbiano caratteristiche intrinseche o ricadute

pratiche del tutto difformi dai principi e valori su cui si fonda la legislazione europea.

Il Regolamento non sarà invece applicabile a settori che non rientrino nell'ambito di applicazione del diritto dell'Unione, non pregiudicando in ogni caso le competenze degli Stati membri in materia di sicurezza nazionale. In sintesi, esso non si applicherà: a) ai sistemi di IA che hanno scopi militari, di difesa o di sicurezza nazionale (indipendentemente dal tipo di entità che svolge tali attività); b) ai sistemi o modelli di IA sviluppati e utilizzati esclusivamente per scopi di ricerca e innovazione; c) ai sistemi rilasciati con licenze free e open source, a meno che non siano collocati sul mercato o immessi in servizio come sistemi ad alto rischio; d) infine, alle persone fisiche che utilizzino l'IA nel corso di attività non professionali, meramente personali.

5. Considerazioni finali: ultimi sviluppi dell'Intelligenza artificiale e prime valutazioni dell'AI Act

L'AI Act ha il grande merito di porsi come prima esperienza, a livello mondiale, nella regolazione complessiva e organica del fenomeno dell'Intelligenza artificiale. Nella prima parte della trattazione si è cercato di evidenziare il significato e la portata della rivoluzione che l'IA di nuova generazione (la c.d. "New Generation AI") sta massicciamente introducendo non solo nella società in cui viviamo – nei suoi diversi profili, come pure in tutte le attività che in essa si svolgono –, ma perfino negli aspetti più privati della nostra vita individuale. Nell'analisi del rapporto tra uomo e computer, prendendo atto della rivoluzione tecnologica in corso, si è evidenziata la necessità di acquisire consapevolezza del fatto che si sono aperti scenari assolutamente nuovi e in gran parte imprevedibili. Il computer non è più la macchina inerte che si limita a eseguire, in modo rigido e sequenziale, gli ordini impartitigli dal programmatore: la "capacità di apprendimento" (*machine learning*) di cui può essere dotato, lo trasforma in interlocutore attivo, in grado non solo di

58. Le definizioni di "fornitore" (provider) e "distributore" (deployer) sono contenute nell'art. 3 dell'AI Act. Unitamente all'individuazione della classe di rischio cui appartiene il sistema di AI volta a volta in esame, l'esatta identificazione di questi due ruoli distinti, ma a volte intersecantisi, proprio per le conseguenze giuridiche che ne derivano, costituisce uno dei problemi centrali nell'applicazione dell'AI Act. Si vedano, tra gli altri: CORAGGIO 2024; CAVAZZUTI 2024; DE LEO-BISCARETTI DI RUFFIA 2024; STEFANELLI-TERRIZZI 2024; GUI-IVANKOVICH-MANCINI et al. 2024; *Deployment*, Voce in "Wikipedia. L'enciclopedia libera".

registrare enormi masse di dati, ma anche di elaborarle in modo da estrarne nuova conoscenza (*data mining*) e, quindi, di conseguire un'esperienza sulla cui base formulare proposte d'azione o, addirittura, prendere autonome decisioni.

Il legislatore europeo, già nel corso della lunga e faticosa elaborazione della sua proposta regolatoria sull'IA, si è trovato a dover affrontare problematiche del tutto nuove o comunque non assimilabili a esperienze normative anche recenti. In particolare, le sempre più numerose e varie applicazioni dell'Intelligenza artificiale generativa (*Generative AI*), che utilizzano algoritmi complessi al fine di creare in modo autonomo (senza l'intervento dell'uomo) nuovi contenuti (non solo testi, ma anche immagini, video, musica), hanno reso necessaria, rispetto all'originaria stesura della Proposta di regolamento da parte della Commissione, l'introduzione d'un complesso insieme di norme in materia di "Modelli di IA per finalità generali" e "Modelli di IA con finalità generali con rischio sistemico" (Regolamento (UE) 2024/1689, Capo V, Sezione 1).

Non si tratta semplicemente dell'utilizzo di nuovi strumenti e di nuove forme d'interazione con la macchina, bensì dell'avvio di una nuova era, contrassegnata da un radicale "cambiamento della nostra architettura cognitiva". Quando si interagisce in maniera non superficiale con sistemi di IA è infatti inevitabile riscontrare mutamenti anche nel proprio modo di affrontare i problemi, sollevando dubbi, formulando (e ponendosi) domande, individuando alternative, analizzando e collegando risposte. Gradatamente, ma irreversibilmente, si sta configurando – e rapidamente estendendo – una nuova forma di pensiero che potrebbe qualificarsi come "ibrida", in quanto in essa si rapportano e si integrano la vastità dei dati che la macchina è in grado d'acquisire ed elaborare, da una parte, e la sensibilità umana, la imprevedibilità

dell'intuizione, la maturazione indotta dall'esperienza, dall'altra⁵⁹.

In considerazione della continua e veloce evoluzione della tecnologia in questo campo, la materia oggetto di regolazione nell'AI Act risulta essersi già ampliata e modificata. Dopo l'ascesa della "IA generativa", infatti, l'IA è sull'orlo di un'altra significativa trasformazione con l'avvento della "IA agentica" (*Agentic AI*). Si tratta di un cambiamento indotto dall'evoluzione dei *Large Language Models* (LLM) in entità attive e decisionali, in grado non solo di generare testi simili a quelli prodotti dall'uomo, ma anche di ragionare, pianificare, usare strumenti ed eseguire autonomamente attività complesse. Questa evoluzione, ridefinendo il modo in cui interagiamo con l'IA, segna l'inizio di una nuova era nell'ambito delle modalità di configurazione e delle possibilità di utilizzo delle corrispondenti tecnologie. L'IA agentica si riferisce a sistemi o agenti che possono svolgere compiti in modo indipendente (senza una guida umana costante), adattandosi anche a situazioni mutevoli. A differenza dei sistemi di IA convenzionali, limitati a compiti fissi, l'IA agentica è dinamica, impara dalle interazioni, migliora progressivamente il suo comportamento nel tempo e ha la capacità di suddividere i compiti in passaggi più semplici, analizzando diverse soluzioni alternative e prendendo decisioni sulla base di obiettivi, istruzioni o *feedback*⁶⁰.

È evidente che il legislatore europeo dovrà prendere sollecitamente in considerazione queste nuove potenzialità emerse nell'ambito dell'IA, in modo da valutarne, sotto i diversi profili giuridici, la varietà delle forme applicative, così come la gravità dei rischi connessi. Rimanendo entro i confini assegnati alla presente trattazione è possibile in questa sede solo accennare a quelle tendenze evolutive che, secondo quanto autorevolmente affermato⁶¹, sono ormai prossime a trasformare

59. In generale, sulle tendenze di sviluppo dell'IA si veda MARTINO 2024. Una panoramica dettagliata delle sfide, delle prospettive e dei rischi legati all'Intelligenza artificiale in vari ambiti (calcolo, big data, gioco, scrittura, etica, tag, lavoro, guerra, cervello, futuro) è presentata nel volume *Intelligenza artificiale. Dieci ambiti della nostra vita che stanno cambiando per sempre*, pubblicato, con la collaborazione degli editor della rivista "Internazionale" D. Cassandro e P. Romano, per BUR da Mondadori, 2024.

60. Sulla "Intelligenza artificiale agentica" si vedano, tra gli altri: ZIA 2024; VILIOTTI 2024; CHUDLEIGH 2024.

61. Si veda STUDIO GARTNER 2024. Sulla "originalità" della conoscenza prodotta dalla IA generativa e sul dialogo come forma di interazione tra l'uomo e la macchina, ma anche come fonte di nuova conoscenza si veda MARTINO 2024.

radicalmente la società, soprattutto sotto il profilo del rapporto uomo-macchina. La GenAI (“AI generativa”) di tipo multimodale (testo, immagine, audio e video), i “Modelli GenAI specifici per il dominio” e gli “Agenti autonomi” sono stati identificati come tecnologie ad alto impatto trasformativo generale, essendo in grado di offrire, tra l’altro, maggiore precisione, trasparenza, sicurezza e privacy⁶². In generale, poi, l’integrazione dell’AI generativa nelle attività di aziende, professionisti e utenti comuni è diventata una tendenza in forte crescita, evidenziando un impatto senza precedenti nella storia delle innovazioni tecnologiche. A ciò si aggiunga ancora che l’industria degli smartphone e dei computer si è rivolta con ottimismo all’Intelligenza artificiale come a un catalizzatore d’innovazione: si prevede che i primi smartphone con AI integrata sui chip miglioreranno radicalmente l’esperienza degli utenti, offrendo assistenti vocali più intelligenti e App in grado d’eseguire compiti complessi.

La sensibilizzazione ai numerosi profili giuridici presentati nell’AI Act dovrebbe indurre a perseguire non solo il potenziamento dei processi di AI, ma anche la loro trasparenza e comprensibilità. In tale contesto si inserisce la c.d. “Intelligenza artificiale spiegabile” (XAI – *Explainable Artificial Intelligence*): si tratta di un’area di ricerca in rapida crescita che si è posta come obiettivo il miglioramento della interpretabilità dei modelli di AI, sovente percepiti come vere e proprie “scatole nere”. Rendere i sistemi di IA non solo precisi e funzionali, ma anche comprensibili per gli utenti genera fiducia e sicurezza, oltre a costituire la necessaria premessa per una documentata valutazione di conformità⁶³.

Sempre con riferimento alle nuove sfide che la regolazione giuridica dell’IA deve affrontare occorre tener presente anche che, molto di recente, si stanno realizzando ulteriori progressi nel senso del superamento della stessa IA generativa: si tratta dell’emergere di sistemi di IA di tipo non solo

“generativo”, ma anche “evolutivo”. Sotto il profilo metodologico un esempio rilevante e concreto di questa transizione verso la “IA evolutiva” è costituito dalla tecnologia “FunSearch” (abbreviazione di “ricerca nello spazio delle funzioni”), sviluppata da Google DeepMind: partendo dalla descrizione di un problema in forma di codice informatico, la procedura ne cerca la soluzione utilizzando la creatività dei *Large Language Models* (LLM) in associazione con un sistema di verifica dell’efficacia delle idee proposte (valutatore sistematico). La particolarità di questa strategia – che rappresenta anche un modello per i futuri sviluppi delle tecnologie nel campo della IA – risiede nell’approccio iterativo, in cui si combinano un’Intelligenza artificiale creativa e una procedura che verifica l’efficacia delle soluzioni proposte⁶⁴.

Notevoli problemi, non solo di natura giuridica, ma anche economica, sociale ed etica, si ricollegano, poi, ai recenti sviluppi della robotica. Progetti come KnowNo⁶⁵, che nella loro configurazione integrano l’Intelligenza artificiale generativa consentendo la realizzazione di robot in grado di “sapere quando non sanno” e quindi di chiedere aiuto solo quando ciò risulti necessario, segnano già l’inizio di una nuova era relativamente al modo in cui queste macchine interagiscono col mondo fisico ed effettuano scelte. Il superamento dei metodi tradizionali di programmazione in questo settore particolare rappresenta un passo significativo verso un futuro in cui i robot non saranno più strumenti passivi, bensì partner affidabili, capaci di operare anche in ambienti dinamici e non strutturati, oltre che di partecipare attivamente nel risolvere problemi e prendere decisioni.

A conclusione di questa rassegna – sia pure superficiale e lacunosa – delle più recenti tendenze evolutive dell’IA, considerate nella prospettiva dei necessari interventi regolatori del legislatore, tra le numerose trasformazioni indotte dalla tecnologia nella struttura sociale e negli ambienti di lavoro occorre accennare anche al fenomeno

62. Una sintetica presentazione delle tecnologie da ultimo citate nel testo è in AI4BUSINESS (Red.) 2024.

63. Un’analisi della *Explainable Artificial Intelligence*, condotta con l’intento di promuovere la discussione collaborativa e la cooperazione interdisciplinare, è in LONGO–BRCIC–CABITZA et al. 2024. Presentazioni più semplici e sintetiche sono in DO COUTO 2024; IBM (Red.) 2020. Un’ampia bibliografia è riportata in *Explainable artificial intelligence*, Voce in “Wikipedia. L’enciclopedia libera”.

64. Si veda FAWZI–ROMERA–PAREDES 2023; ROMERA–PAREDES–BAREKATAIN–NOVIKOV et al. 2024; MOURET 2024.

65. Si veda REN–DIXIT–BODROVA et al. 2023; VILIOTTI 2023.

conosciuto come BYOAI (acronimo di *Bring Your Own Artificial Intelligence*). In uno scenario del genere i dipendenti di un'azienda, per svolgere i compiti loro assegnati, utilizzano strumenti e applicazioni di IA esterni al proprio contesto lavorativo, indipendentemente dal fatto che tali strumenti siano o meno ufficialmente approvati dall'azienda stessa. Com'è agevole intuire, questa tendenza, alimentata dalla crescente disponibilità di sistemi di IA a basso costo e di semplice utilizzo, fa sorgere notevoli problemi giuridici riguardanti soprattutto la governance organizzativa e la valutazione della conformità normativa, ma anche il controllo e la sicurezza dei dati, il rispetto del copyright e la tutela della privacy⁶⁶.

L'AI Act, come già rilevato, costituisce il primo tentativo, a livello mondiale, di regolazione normativa dell'Intelligenza artificiale e delle sue estremamente numerose e differenti applicazioni. Nello svolgimento della presente trattazione si è cercato di mettere in luce le caratteristiche essenziali e lo sconvolgente potenziale trasformativo di questo fenomeno del tutto nuovo nella storia dell'intera umanità. Si è reso evidente che, da un lato, i molteplici problemi sollevati richiedono interventi tempestivi; dall'altro, proprio in funzione dei delicati e complessi equilibri che si rischia di compromettere, le iniziative regolatorie necessitano di un'attenta ponderazione, ma anche di una costante rivalutazione.

Anche a un livello d'analisi superficiale, in prima istanza risulta assai arduo definire i confini della materia da regolare, in considerazione dell'attuale sempre più spinta pervasività dell'IA, ma anche dell'impressionante molteplicità e varietà delle sue applicazioni. Ogni tipologia d'utilizzo – spaziando senza limiti (ed esemplificando in maniera del tutto casuale) dalla medicina alla finanza, dalla ricerca scientifica all'organizzazione aziendale, dalla mobilità all'assistenza legale – coinvolge numerose specializzazioni, sovente distanti tra loro e non sempre agevolmente reperibili. Non solo; ad ogni contesto applicativo corrispondono problematiche

differenti, di tipo sociale, giuridico ed etico. A tutto ciò si aggiunge il fatto che se, da una parte, appare estremamente complesso individuare con esattezza l'oggetto della normazione, dall'altra, sotto il profilo temporale il legislatore si trova nell'impossibilità di prevedere e fronteggiare gli sviluppi della tecnologia, in considerazione dell'impressionante velocità del progresso scientifico del settore, ma anche del continuo superamento di limiti che solo poco tempo prima sembravano invalicabili.

Tra l'altro, lo sviluppo esplosivo dell'IA ha incluso anche un processo di "democratizzazione", coinciso con il lancio, nel 2023, di "ChatGPT", un chat-bot AI che, sfruttando l'Intelligenza artificiale generativa e l'apprendimento automatico e basandosi su un modello conversazionale, è in grado di dialogare con l'utente finale e creare non solo testi, ma anche immagini, audio e video⁶⁷. A livello mondiale migliaia di sistemi di IA fondati sulle tecnologie di OpenAI, Google AI e Microsoft si stanno rapidamente diffondendo. Inoltre, Google ora integra la propria Intelligenza artificiale in tutte le App e, allo stesso modo, Microsoft utilizza la sua IA in tutti i programmi, compresi quelli che fanno parte di Microsoft Office. Anche molte terze parti, tra cui importanti reti social come LinkedIn, stanno integrando l'IA nelle loro piattaforme. Occorre rilevare, però, che, rispetto a questa realtà, sembra mancare, o essere comunque carente e non aggiornata, la competenza dei politici e dei legislatori, nei Parlamenti nazionali come a livello europeo.

Probabilmente proprio a causa delle difficoltà indicate, derivanti dai labili confini dell'Intelligenza artificiale e dalla rapidità della sua evoluzione, l'AI Act si presenta come un documento molto lungo, complesso e sovente di difficile interpretazione. Il testo normativo comprende 180 considerando, 113 articoli (organizzati in 13 Capi e talora tanto estesi da dover essere suddivisi in paragrafi, commi, numeri, lettere e punti⁶⁸) e 13 Allegati. Inoltre, sono molto numerose le disposizioni di rilievo sostanziale che dovranno essere definite solo in atti legislativi e regolamentari successivi, la

66. Si vedano VAN DER MEULEN–WIXOM 2024, MARR 2024.

67. Molto popolare in Italia fin dal suo debutto, ChatGPT ha riscosso ampi consensi tra gli italiani, che risultano tra i primi 10 utenti al mondo per volume di ricerche. L'ultima versione, denominata GPT-4o, è stata presentata il 13 maggio 2024. Si veda CASTIGLI–SANDONNINI 2024.

68. Si vedano, ad esempio, l'art. 3, contenente ben 68 definizioni esplicative di espressioni e termini tecnici utilizzati nella legge, e l'art. 5, che enumera le "Pratiche di IA vietate".

cui emanazione è posta a carico della Commissione⁶⁹. A questi occorre poi aggiungere le norme con cui i 27 Stati membri recepiranno il Regolamento. Il rinvio ad atti successivi può apparire come un'opportuna scelta di flessibilità, uno strumento idoneo a fronteggiare una tecnologia in continua e rapida evoluzione, ma rende comunque manifesta la difficoltà d'imbrigliare una materia assai vasta e complessa in un unico atto normativo. Ovviamente, le modalità con cui la Commissione e gli Stati membri applicheranno in concreto le disposizioni del Regolamento andranno poi a incidere sulla valutazione complessiva dell'efficacia della legge.

Sul piano pratico, però, il problema di fondo, uniformemente evidenziato da studiosi e ricercatori, come da imprenditori, economisti e politici, risiede nella situazione di grave ritardo in cui l'Europa versa nei confronti dei principali *competitor* mondiali nel campo dell'Intelligenza artificiale. Nel decennio 2013-2023 risultano oltre 5.500 società specializzate in IA fondate negli Stati Uniti: si tratta d'un numero 4 volte superiore alla Cina, ma addirittura superiore rispetto a quello di tutti gli altri Paesi messi insieme⁷⁰. In Europa ci sono tuttora pochissime aziende in questo settore e mancano, o sono del tutto insufficienti, alcune delle infrastrutture e delle componenti critiche fondamentali (*cloud computing*, *cybersecurity*, produzione di microchip). In un contesto del genere l'AI Act potrebbe rappresentare il tentativo, da parte dell'Europa, di condizionare mediante una rigida normativa la corsa di Stati Uniti e Cina, nell'impossibilità di concorrere con loro sul piano degli investimenti e dello sviluppo tecnologico. È essenziale, invece, che l'Europa scenda decisamente in campo non solo come arbitro di una competizione tra le due superpotenze, ma soprattutto come protagonista di una sfida irrinunciabile, nella concreta prospettiva dell'acquisizione, in un futuro non lontano, di una leadership globale: aumentare

gli investimenti pubblici e privati in ricerca e innovazione all'interno di un efficiente coordinamento europeo dovrebbe portare a creare, in tempi rapidi, la massa critica necessaria a consentire una partecipazione attiva al dialogo e al confronto.

Le critiche prontamente mosse da più parti all'AI Act riguardano molti aspetti della vasta problematica emergente, spaziando, per citarne solo alcune, dai contrasti rilevabili con le precedenti normative europee di settore alla vaghezza di certe definizioni, dalle difficoltà – soprattutto per le piccole e medie imprese – di rispettare meccanismi regolatori costosi e complessi alla pesante incidenza del sistema sanzionatorio, dalla possibilità di esercitare un controllo effettivo sul contenuto dei dataset utilizzati per l'addestramento degli algoritmi alle insidie nascoste nella capacità di "allineamento" delle risposte del programma alle domande e alle attese dell'utente. Le accuse più pesanti e coinvolgenti riguardano, però, proprio la materia dei diritti fondamentali. La nuova legge sull'IA è giudicata da alcuni piena di concessioni alle lobby della grande industria del settore⁷¹, avendo accolto esenzioni relative agli usi più pericolosi dell'IA soprattutto da parte delle forze dell'ordine e delle autorità per la migrazione, e contenendo divieti che, in quanto includono numerose eccezioni e scappatoie, in realtà finirebbero col non vietare alcuni degli usi più pericolosi delle nuove tecnologie, come quello relativo ai sistemi d'identificazione biometrica remota in tempo reale, in spazi accessibili al pubblico e a fini di attività di contrasto⁷².

Queste sintetiche osservazioni evidenziano il problema centrale di fronte al quale si è trovato il legislatore europeo: si trattava, da una parte, di mettere l'Unione in condizioni di svolgere un ruolo da protagonista nella scena mondiale della rapidissima trasformazione tecnologica in atto, dall'altra, di garantire comunque la tutela di quei valori e principi su cui l'Unione stessa si fonda.

69. È stato calcolato che le espressioni "*delegated act*" (decreti delegati, che richiedono l'approvazione del Parlamento) e "*implementing act*" (regolamenti, che non richiedono la successiva approvazione del Parlamento) compaiono, rispettivamente, ben 34 e 39 volte. Si veda CAPACCI-GALLI-LOREGGIA-MAROCCIA 2024.

70. Si veda D'ASCENZO 2024.

71. Per una sintetica rassegna delle principali criticità evidenziate nell'AI Act, si vedano, tra gli altri: ACCESS NOW (Red.) 2023; BENDER 2023; VRANKEN 2023; ALFIERI 2024; BONAVITA-CORTINA 2024; CIRONE 2024; DE BREE 2024; PANETTA 2024; PONTI 2024; SANDONNINI 2024; SPERONI DI FENIZIO 2024; MELE 2025.

72. Si veda, in proposito, all'interno dell'AI Act, l'art. 5, par. 1, primo comma, lettera h e, più in generale, la fitta rete di rinvii interni ed esterni che integrano il complesso meccanismo di limitazioni e controlli.

Quanto il difficile contemperamento di queste distinte esigenze sia di fatto riuscito deve essere certamente oggetto d'analisi approfondite, il più possibile esenti da pregiudizi e condizionamenti ideologici. In chiusura ci si limita qui a richiamare le direttive indicate dall'UNESCO nella sua Raccomandazione sull'etica dell'Intelligenza artificiale, volta a fornire standard globali per massimizzare i benefici dell'innovazione tecnologica, assicurando che essa contribuisca realmente a costruire un mondo più inclusivo, sostenibile e pacifico⁷³.

Riferimenti bibliografici

- S.A. AARONSON (2020), *America's uneven approach to AI and its consequences*, in "Institute for International Economic Policy Working Paper Series", George Washington University, April 2020
- ACCESS NOW (Red.) (2023), *The EU should regulate AI on the basis of rights, not risks*, in "Access Now", 17 February 2021 (last updated: 13 January 2023)
- AI4BUSINESS (Red.) (2024), *Gartner: entro il 2027 il 40% delle soluzioni di AI generativa sarà multimodale*, in "AI4Business", 9 settembre 2024
- A. ALAIMO (2023), *Il Regolamento sull'Intelligenza Artificiale: dalla proposta della Commissione al testo approvato dal Parlamento. Ha ancora senso il pensiero pessimistico?*, in "federalismi.it", 2023, n. 25
- G. ALFIERI (2024), *Intelligenza artificiale, chi critica e perché l'AI Act Ue*, in "Startmag", 14 marzo 2024
- V. ALLIS (1994), *Searching for Solutions in Games and Artificial Intelligence*, Ph.D. Thesis, Maastricht University, 1994
- N. ANAND (2024), *US Federal AI Legislation in 2024: The Current Landscape*, in "Holistic AI", March 8 2024
- M. BARBAJA (2024), *Il quadro normativo UE sull'Intelligenza Artificiale: gli aspetti innovativi del nuovo regolamento*, in "De Iustitia", n. 3, 2024
- G. BAUM (2023), *La commedia ai tempi di ChatGPT. Ovvero, ecco perché l'IA non è divertente (almeno per ora)*, in "The Hollywood Reporter", 6 giugno 2023
- G. BENDER (2023), *The global AI race – it's time to slow down*, in "Social Europe", 15 February 2023
- J.R. BIDEN (2023), *Safe, Secure and Trustworthy Development and Use of Artificial Intelligence*, Executive Order 14110 of October 30, 2023
- E. BOCCHINI (2024), *La regolazione giuridica dell'intelligenza artificiale*, Giappichelli, 2024
- S. BONAVITA, A. CORTINA (2024), *L'AI Act a rischio corto circuito: ecco i problemi nell'applicazione pratica*, in "Agenda digitale", 2 ottobre 2024

73. Sui rischi che i sistemi di IA possono presentare nei confronti dei diritti fondamentali e sulla conseguente necessità di elaborare "protezioni etiche" ("ethical guardrails") a livello mondiale si veda: United Nations Educational, Scientific and Cultural Organization – UNESCO, Recommendation on the Ethics of Artificial Intelligence: key facts, (adopted on 23 November 2021), 2023. Il [messaggio introduttivo](#) di Gabriela Ramos (Social and Human Sciences, UNESCO) è sul sito dell'UNESCO. Un'altra importante iniziativa promossa dall'UNESCO in questo campo è costituita dal Global AI Ethics and Governance Observatory, Osservatorio realizzato con l'obiettivo di fornire una risorsa globale ai decisori politici, ma anche agli accademici, al settore privato e alla società civile in genere, al fine di elaborare soluzioni alle sfide più urgenti sollevate dall'Intelligenza artificiale. L'Osservatorio ospita, a sua volta, l'AI Ethics and Governance Lab, che raccoglie documentazione relativa a casi di studio, buone pratiche e ricerche di punta, in modo da fornire un supporto concreto e strumenti innovativi per un adeguato inquadramento delle principali problematiche connesse all'IA.

- A. BURT (2021), *New AI Regulations Are Coming. Is Your Organization Ready?*, in “Harvard Business Review”, April 30, 2021
- L. CAMAIONI (2016), *La Teoria della mente. Origini, sviluppo e patologia*, Fabbri, 2016
- A. CAPACCI, G. GALLI, A. LOREGGIA, I. MAROCCIA (2024), *Il nuovo regolamento europeo sull’IA: cosa cerca di fare e cosa fa*, in “Osservatorio CPI, Conti Pubblici Italiani”, 22 febbraio 2024
- M. CASTIGLI, P. SANDONNINI (2024), *ChatGPT guida completa: cos’è, come si usa e cosa può fare per aziende e professionisti*, in “Agenda digitale”, 25 ottobre 2024
- C. CATH, S. WATCHER, B. MITTELSTADT et al. (2018), *Artificial intelligence and the “Good Society”: the US, EU and UK approach*, in “Science and Engineering Ethics”, vol. 24, 2018
- A. CAVAZZUTI (2024), *Provider e Deployer nell’AI Act: chi sono e cosa fanno?*, in “Studio legale Delli Ponti”, 16 aprile 2024
- E. CHITI, B. MARCHETTI (2020), *Divergenti? Le strategie di Unione europea e Stati Uniti in materia di intelligenza artificiale*, in “Rivista della regolazione dei mercati”, 2020, n. 1
- S. CHUDLEIGH (2024), *Che cos’è un agente AI*, in “botpress”, 31 maggio 2024
- E. CIRONE (2024), *L’AI Act e l’obiettivo (mancato?) di promuovere uno standard globale per la tutela dei diritti fondamentali*, in “Quaderni AISDUE”, n. 2/2024, fascicolo speciale su “La nuova disciplina UE sull’intelligenza artificiale”
- COMMISSIONE EUROPEA (2021), *Piano coordinato sull’intelligenza artificiale*, 2021
- COMMISSIONE EUROPEA (2020), *Un nuovo piano d’azione per l’economia circolare. Per un’Europa più pulita e più competitiva*, COM(2020) 98, 11 marzo 2020
- CONSIGLIO DI STATO CINESE (2017), *New Generation Artificial Intelligence Development Plan – AIDP* (trad. a cura di G. Webster, R. Creemers, E. Kania, P. Triolo), in “DigiChina”, Stanford University, August 1, 2017
- CONSIGLIO EUROPEO, CONSIGLIO DELL’UNIONE EUROPEA (2024), *In che modo l’UE rafforza la propria cibern sicurezza*, 2 dicembre 2024
- G. CORAGGIO (2024), *Siete un Provider o un Deployer ai sensi dell’AI Act?*, in “Diritto al digitale”, 2 aprile 2024
- CORTE DEI CONTI EUROPEA (2024), *Intelligenza artificiale: l’UE deve accelerare il passo*, 29 maggio 2024
- M. D’ASCENZO (2024), *Venture capital, in Europa i fondi puntano sull’intelligenza artificiale*, in “Il Sole 24 ore”, 18 ottobre 2024
- T. DE BREE (2024), *La scomoda verità della normativa europea sull’intelligenza artificiale (AI Act)*, in “Risk & Compliance”, 5 gennaio 2024
- M. DE LEO, B. BISCARETTI DI RUFFIA (2024), *La responsabilità del deployer nell’uso dell’Intelligenza Artificiale*, in “DB - Non solo diritto bancario”, 17 dicembre 2024
- M. DO COUTO (2024), *Entering the Age of Explainable AI*, in “TDWI – Transforming Data With Intelligence”, 2024
- DOUGLAS HEAVEN (2017), *The epic robot fails that say ai will never rule the world. Will the machines rise up and enslave us? The history of incompetent robots, from klutzy cleaners to sense-mangling chatbots, suggests we can sleep easy*, in “New Scientist”, 19 December 2017
- E. DUSI (2024), *Il Premio Nobel per la Fisica a John Hopfield e Geoffrey Hinton per l’intelligenza artificiale*, in “La Repubblica”, 8 ottobre 2024

- F. FAGGIN (2024), *Oltre l'invisibile. Dove scienza e spiritualità si uniscono*, Mondadori, 2024
- F. FAGGIN (2022), *Irriducibile. La coscienza, la vita, i computer e la nostra natura*, Mondadori, 2022
- A. FAWZI, B. ROMERA-PAREDES (2023), *FunSearch: Making new discoveries in mathematical sciences using Large Language Models*, in "Google DeepMind", 14 December 2023
- E. FLEISCHNER (2023), *Perché non bisogna avere paura dell'Intelligenza Artificiale, ma solo fare attenzione*, in "Spirito Artigiano", 16 settembre 2023
- GARANTE PER LA PROTEZIONE DEI DATI PERSONALI (2021), *Assistenti digitali (smart assistant): i consigli del Garante per un uso a prova di privacy, Scheda informativa*, marzo 2021
- A. GLUSHENKOV (2024), *Deep Blue: The Chess Supercomputer That Changed AI and IBM Forever*, in "Medium", November 2, 2024
- A. GUI, P. IVANKOVICH, S. MANCINI et al. (2024), *AI Act, come definire ruoli e responsabilità per evitare sanzioni: la guida per imprese e P.A.*, in "Agenda digitale", 4 novembre 2024
- F. HSU (2004), *Behind Deep Blue: Building the Computer That Defeated the World Chess Champion*, Princeton University Press, 2004
- M. IASELLI (a cura di) (2024), *AI Act. Principi, regole ed applicazioni pratiche del Reg. UE 1689/2024*, Maggioli editore, 2024
- IBM (Red.) (2020), *Che cos'è l'AI spiegabile (XAI)*, 2020
- E. INTINI (2024), *Geoffrey Hinton, quella volta che il Nobel "padrino pentito dell'AI" lasciò Google*, in "Focus", 8 ottobre 2024
- E. INTINI (2017), *Hawking: Il nostro tempo come specie dominante sulla Terra sta per scadere*, in "Focus", 9 novembre 2017
- S. KHANAL, H. ZHANG, A. TAEIHAGH (2024), *Development of New Generation of Artificial Intelligence in China: When Beijing's Global Ambitions Meet Local Realities*, in "Journal of Contemporary China", 9 April 2024
- V. LALLI (2024), *AI Act: approvato il testo del Regolamento UE sulla Intelligenza Artificiale*, in "Avvocloud.net. La rete di avvocati online", 29 settembre 2024
- J. LANDGREBE, B. SMITH (2022), *Why Machines Will Never Rule the World. Artificial Intelligence without Fear*, Routledge, 2022
- S. LECCE, E. CAVALLINI, A. PAGNIN (2010), *La teoria della mente nell'arco di vita*, il Mulino, 2010
- L. LONGO, M. BRCIC, F. CABITZA et al. (2024), *Explainable Artificial Intelligence (XAI) 2.0: A Manifest of Open Challenges and Interdisciplinary Research Directions*, in "Information Fusion", vol. 106, June 2024
- J. MA (2024), *The U.S. reigns supreme in A.I. startups while China ensures chatbots have 'core socialist values'*, in "Fortune", July 21, 2024
- S. MAGGIOLO (2016), *L'intelligenza di AlphaGo*, in "Il Post", 21 marzo 2016
- B. MARR (2024), *7 Ways To Turn The "Bring Your Own AI" Threat Into An Opportunity*, in "Forbes", August 22, 2024
- A.A. MARTINO (2024), *Dove sta andando l'intelligenza artificiale?*, in "Il Sole 24 ore. NT+Diritto", 4 ottobre 2024
- A. MCAFEE, E. BRYNJOLFSSON (2016), *Where Computers Defeat Humans and Where They Can't*, in "The New York Times", March 16, 2016

- S. MELE (2025), *Rischi e opportunità dell'AI Act in attesa del Vertice di Parigi*, in "Formiche", 8 febbraio 2025
- R. METZ (2018), *Microsoft's neo-Nazi sexbot was a great lesson for makers of AI assistants*, in "MIT Technology Review", March 27, 2018
- M. MIALICH (2021), *Smart Assistant: profili di criticità e raccomandazioni delle Autorità per la protezione dei dati personali*, in "CyberLaws", 21 giugno 2021
- J. MOURET (2024), *Large Language Models help computer programs to evolve*, in "Nature", 17 January 2024
- E. MUSK et al. (2017), *An Open Letter to the United Nations Convention on Certain Conventional Weapons*, Future of Life Institute, 20 August, 2017
- NEW SCIENTIST (2018), *Machines that Think. Everything you need to know about the coming age of artificial intelligence* (trad. it. V.L. Gili, "Macchine che pensano. La nuova era dell'Intelligenza Artificiale", Edizioni Dedalo, dicembre 2018)
- R. PANETTA (2024), *AI Act: critiche sbagliate, era urgente regolare l'Intelligenza Artificiale*, in "Agenda digitale", 26 marzo 2024
- PARLAMENTO EUROPEO (2023), *Economia circolare: definizione, importanza e vantaggi*, 25 maggio 2023
- M. PEDRAZZOLI (2024), *La Corte dei conti UE lancia l'allarme: nell'Unione non si sta facendo abbastanza per sviluppare l'Intelligenza artificiale*, in "EU News", 29 maggio 2024
- M. POLANYI (1966/1979), *The Tacit Dimension*, 1966 (trad. it. "La conoscenza inespresa", Armando Editore, 1979)
- M. POLANYI (1958/1990), *Personal Knowledge: Towards a Post-Critical Philosophy*, 1958 (tr. it. "La conoscenza personale", Rusconi, 1990)
- C. PONTI (2024), *AI Act, sacrificati i diritti dei più deboli: ecco perché*, in "Cybersecurity 360", 15 marzo 2024
- A.Z. REN, A. DIXIT, A. BODROVA et al. (2023), *Robots That Ask For Help: Uncertainty Alignment for Large Language Model Planners*, in "arXiv", 4 September 2023
- LA REPUBBLICA (Red.) (2025), *Lee SeDol giocava a Go come un Dio. Ma da quando una IA l'ha battuto, non è più lo stesso*, 16 luglio 2024 (aggiornato 20 febbraio 2025)
- LA REPUBBLICA (Red.) (2016), *Microsoft, il chat-bot Tay scrive frasi razziste: cancellati i suoi tweet*, 25 marzo 2016
- H. ROBERTS, J. COWLS, J. MORLEY et al. (2021), *The Chinese Approach to AI: An Analysis of Policy, Ethics, and Regulation*, in "AI and Society", vol. 36, 2021, n. 1
- B. ROMERA-PAREDES, M. BAREKATAIN, A. NOVIKOV et al. (2024), *Mathematical discoveries from program search with large language models*, in "Nature", vol. 625, 2024
- M. ROVELLI, P. OTTOLINA (2024), *Chi è Geoffrey Hinton, «padrino» pentito dell'intelligenza artificiale, che ha vinto il Nobel: «È spaventosa»*, in "Corriere della sera", 8 ottobre 2024
- P. SANDONNINI (2024), *AI Act in vigore dal 1° agosto, ma già piovono critiche. E l'Ue si difende*, in "AI4Business", 16 luglio 2024
- C. SHANNON (1950), *Programming a Computer for Playing Chess*, in "Philosophical Magazine", vol. 41, 1950, n. 314
- D. SILVER, A. HUANG, C.J. MADDISON et al. (2016), *Mastering the Game of Go with Deep Neural Networks and Tree Search*, in "Nature", vol. 529, 27 January 2016

- P. SPERONI DI FENIZIO (2024), *L'Eu AI Act e i suoi problemi*, in "Futura Network", 11 luglio 2024
- S. STEFANELLI, A. TERRIZZI (2024), *AI Act e responsabilità penale, cosa cambia per provider e deployer*, in "Agenda digitale", 17 luglio 2024
- STUDIO GARTNER (2024), *Le 10 principali tendenze tecnologiche strategiche per il 2025*, in "AI4Business", 22 ottobre 2024
- S. SULMICELLI (2023), *Stati Uniti d'America – Executive Order on The Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, in "BioDiritto", 30 ottobre 2023
- U.S. DEPARTMENT OF COMMERCE, OFFICE OF PUBLIC AFFAIRS (2024), *U.S. and UK Announce Partnership on Science of AI Safety*, April 1, 2024
- N. VAN DER MEULEN, B. H. WIXOM (2024), *Bring Your Own AI: How to Balance Risks and Innovation*, in "MIT Sloan Management Review", October 3, 2024
- A. VILIOTTI (2024), *Agenti AI: Intelligenza Artificiale per il pensiero, l'organizzazione e l'impiego di strumenti*, 29 aprile 2024
- A. VILIOTTI (2023), *Scopri come KnowNo sta cambiando il futuro della robotica*, in "Linkedin", 13 dicembre 2023
- P. VOIGT, N. HULLEN (2024), *The EU AI Act: Answers to Frequently Asked Questions*, Springer, 2024
- B. VRANKEN (2023), *Big Tech lobbying is derailing the AI Act*, in "Corporate Europe Observatory", 24 November 2023
- M.J. WOLF, K.W. MILLER, F.S. GRODZINSKY (2017), *Why We Should Have Seen That Coming. Comments on Microsoft's Tay "Experiment", and Wider Implications*, in "The ORBIT Journal", vol. 1, 2017, n. 2
- T. ZIA (2024), *Agentic AI: come i grandi modelli linguistici stanno plasmando il futuro degli agenti autonomi*, in "Unite.AI", 1° novembre 2024
- L. ZORLONI (2024), *Inizia la stagione dell'AI Act: ecco le 6 scadenze del Regolamento europeo sull'Intelligenza artificiale*, in "Wired", 12 luglio 2024
- R. ZREICK (2018), *Perché non devi avere paura dell'Intelligenza Artificiale*, in "Focus", 21 gennaio 2018
- V. ZUCCONI (1997), *Kasparov – Deep Blue, sfida truccata*, in "la Repubblica", 8 maggio 1997