Nuovi rischi

Intelligenza artificiale e sicurezza sul lavoro: rischi e opportunità

Maria Elena Iafolla - Avvocato, Studio Legale Rosiello e Associati

Che cos'è l'intelligenza artificiale?

Nell'approcciarsi all'intelligenza artificiale (IA), il primo necessario passo deve essere quello di inquadrarla in una definizione, attività questa non semplice, prestandosi essa, per sua natura, a molteplici interpretazioni e alle più diverse applicazioni. L'IA è una disciplina moderna - seppur non quanto ci si aspetterebbe - appartenente all'informatica, ma influenzata da numerose altre materie, quali la filosofia, la matematica, le neuroscienze, la psicologia, le scienze cognitive e la linguistica.
Essa studia i fondamenti teorici, le metodologie e le tecniche attraverso cui progettare sistemi hardware e software capaci di fornire all'elaboratore elettronico delle prestazioni che a un osservatore comune sembrerebbero di pertinenza esclusiva dell'intelligenza umana (1). Obiettivo della materia non è quello di ricostruire in qualche modo l'intelligenza umana, ma piuttosto di emularla, anche con meccanismi differenti da quelli propri dell'uomo, in ogni caso capaci di fornire prestazioni qualitativamente e quantitativamente almeno equivalenti a quelle umane. I modelli utilizzati possono, dunque, essere sia di tipo antropomorfico (in applicazione degli stessi metodi concettuali utilizzati dall'uomo) sia di tipo non antropomorfico, al fine di ottenere i migliori risultati possibili.
Come è stato osservato, inoltre, l'IA è al tempo stesso una scienza e un'ingegneria (2):
— una scienza perché, nel momento in cui vengono emulati, con determinati sistemi artificiali, alcuni dei comportamenti intelligenti (nel senso di riconducibili all'uomo) si realizza uno schema di sperimentazione e creazione di modelli, tipica appunto della scienza;
— parallelamente, è però anche un'ingegneria, poiché permette di ottenere dalle macchine prestazioni "artificiali" che emulano determinati comportamenti.
Vale forse la pena di riportare alcune delle definizioni storicamente più rilevanti e interessanti, anche per ricordare che l'IA è una disciplina estremamente attuale, ma non neonata. A ben guardare, le definizioni risultano interessanti soprattutto per la difficoltà di indicare che cosa voglia dire essere intelligente.
John McCarthy, da molti considerato il vero padre dell'IA, affermò ad esempio che: "il problema dell'intelligenza artificiale consiste nel creare una macchina che si comporti in modi che sarebbero considerati intelligenti se un umano si comportasse nello stesso modo" (3). La definizione fu inserita nella proposta di un progetto di ricerca presso il Dartmouth College del 31 agosto 1955, data che si considera formale nascita della disciplina.
La definizione è stata successivamente ripresa da Marvin Minsky, che nel 1968 parlò di IA come della "scienza di creare macchine che fanno cose che richiederebbero intelligenza se fatte dall'uomo" (4).
L'impostazione salta a piè pari il punto focale e cioè la definizione di che cosa sia l'intelligenza; questo problema - informatico, linguistico, filosofico - viene evidenziato dallo stesso McCarthy, che più di 50 anni dopo, nel 2007, scrisse: "Il problema è che ancora non possiamo caratterizzare in generale quali tipologie di procedure computazionali vogliamo chiamare intelligenti. Noi comprendiamo alcuni dei meccanismi dell'intelligenza, ma non altri" (5).

(1) Somalvico, *L'Intelligenza Artificiale*, Rusconi Editore, Milano, 1987.
(2) Nilsson, *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann, San Mateo, CA, USA, 1998.
(3) "The artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving", in McCarthy, *A proposal for the Dartmouth summer research project on artificial intelligence*, http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html, link consultato il 12 agosto 2021.
(4) Minsky, *Semantic Information Processing*, 1968, The MIT Press, Cambridge, Mass.
(5) McCarthy, *What is artificial intelligence?*, Computer Science Department, Stanford University, 2007 Nov 12, 2:05 a.m. Un metodo molto conosciuto per la verifica dell'IA è il cosiddetto Test di Turing, ideato da Alan Mathison Turing e reso noto

Le applicazioni dell'intelligenza artificiale

Le applicazioni dei sistemi di IA sono moltissime e di difficile catalogazione o anche soltanto elencazione, per la quantità ed eterogeneità dei settori interessati e anche per la continua innovazione che è caratteristica propria della disciplina.
Tra i diversi schemi, i più accreditati ricomprendono *machine learning*, percezione, elaborazione del linguaggio naturale e robotica.

Machine learning

Quest'area della disciplina studia gli algoritmi che imparano e si adattano in base ai dati che ricevono. Per usare un esempio ormai familiare ai più, si pensi ai filtri *anti-spam*, addestrati a riconoscere i messaggi di posta indesiderata, che continuano ad aggiornarsi e ad "imparare" anche una volta in funzione sulla casella di posta.

Percezione

All'interno di quest'ambito si trovano due aree tecnicamente molto diverse tra loro e cioè la visione artificiale e il riconoscimento vocale. La prima è sfruttata soprattutto per classificare e riconoscere oggetti, verificare traiettorie e spazi, identificare persone; il secondo, invece, lavora su un *input* audio, traducendolo in testo, identificando una determinata persona o anche suggerendo la presenza di eventuali patologie, secondo alcune recenti applicazioni alla scienza medica.

Elaborazione del linguaggio naturale

Quest'area (NLP - *natural language processing*) lavora sulla comprensione e quindi sulla produzione del linguaggio umano. Si tratta, dunque, dell'area con cui più facilmente i non addetti ai lavori vengono in contatto, poiché il suo obiettivo principe è proprio quello di premettere l'interazione uomo-macchina, allargando così la platea di fruitori anche a soggetti privi delle necessarie competenze tecniche.

Robotica (avanzata)

La robotica costituisce una disciplina autonoma, ma integra molte funzioni di IA per migliorare le prestazioni delle macchine, potendo il software impartire alla macchina istruzioni su cosa fare e come farlo.
Questi schemi possono avere applicazione nelle più diverse aree, quali la produzione industriale, i giochi, la mobilità, la dimostrazione di teoremi matematici o la programmazione automatica, elenco che non può essere esaustivo perché in continua espansione.

Intelligenza artificiale e sicurezza sul lavoro

Il tema dell'intelligenza artificiale è quanto mai attuale e di grande impatto anche nel mondo della sicurezza sul lavoro, ove essa è già ampiamente utilizzata con una vasta gamma di applicazioni sia per l'analisi di dati e processi gestionali, sia per l'operatività. Ci si riferisce ad esempio alle tecnologie indossabili, all'assistenza nelle linee di assemblaggio e all'assistenza virtuale, all'utilizzo dei DPI o ai *cobot*, i *robot* collaborativi.
Non sarà sfuggito a chi si interessa della materia che il dibattito in tema di IA riguarda per lo più la quantità di posti di lavoro e dunque il rischio di una sensibile riduzione per la possibilità di impiegare macchine di intelligenza artificiale. Non meno interessante, tuttavia, è la riflessione circa i temi di salute e sicurezza sul lavoro, che evidenzia indubbi vantaggi, ma anche rischi da tenere in debita considerazione.
Tra i vantaggi, si pensi, per cominciare, ai *cobot* e cioè a quei *robot* concepiti per interagire fisicamente con l'uomo in uno spazio di lavoro (6). Simili nell'aspetto e nel funzionamento rispetto ai *robot* industriali che abbiamo ormai imparato a conoscere da molto tempo, i *cobot* costituiscono una realtà piuttosto recente nel mondo della produttività, se si pensa che il primo brevetto è stato depositato dalla Northwestern University nel 1997 (7).
La particolarità dei *cobot*, che ne costituisce al contempo un vantaggio e un fattore di rischio, sta nel fatto che essi non devono necessariamente essere programmati per svolgere un compito, ma possono

nell'articolo *Computing Machinery and Intelligence*, 1050, Mind. Si tratta di un esperimento consistente in un "gioco dell'imitazione" a tre partecipanti: A, B e C. C è tenuto separato dagli altri due e può solo stabilire attraverso una serie di domande quale sia l'uomo e quale la donna. Dal canto loro, A dovrà ingannare C, portandolo ad una identificazione errata, mentre B dovrà aiutarlo a mantenere un'identificazione corretta. Il test di Turing si basa sul presupposto che una macchina si sostituisca ad A e, in tal caso, se C non si accorgesse di nulla, la macchina dovrebbe essere considerata intelligente, dal momento che sarebbe indistinguibile da un essere umano.

(6) Una raccolta di approfondimenti anche tecnici è disponibile sul sito della Northwestern University, USA, a questo link: https://peshkin.mech.northwestern.edu/cobot/, consultato il 18 agosto 2021.

(7) Una sintesi degli elementi tecnici è disponibile a questo link: https://patents.google.com/patent/US5952796A/en, consultato il 18 agosto 2021.

adattarsi e "imparare" la sequenza di azioni da compiere, secondo la metodologia del *machine learning* sopra introdotta. L'altra caratteristica, che a sua volta rappresenta un vantaggio e un fattore di rischio, sta nella condivisione dello spazio tra uomo e macchina, il cosiddetto *collaborative workspace*.
La robotica collaborativa permette di affidare a macchine veloci, accurate e instancabili i compiti ripetitivi e le attività pericolose o svolte in luoghi pericolosi e permette d'altra parte di agevolare l'accesso al lavoro di molte persone che ne restano potenzialmente escluse, per esempio aiutando i disabili o i lavoratori meno giovani sul luogo di lavoro.
Parallelamente, esistono alcuni fattori di rischio per la salute e sicurezza dei lavoratori che non vanno dimenticati: si pensi, ad esempio, alle ipotesi in cui uomo e macchina debbano collaborare a contatto e dunque sussista la possibilità di urti e incidenti o, addirittura, di manomissioni dolose della macchina connessa ad *internet*.
Essenziali sono, dunque, attività tecniche come il costante monitoraggio delle impostazioni di programmazione (e delle attività apprese dalla macchina in *machine learning*) e una elevata attenzione alla sicurezza informatica; parallelamente, misura imprescindibile resta la formazione dei lavoratori, con aggiornamenti periodici e mirate campagne formative e informative.
Accanto ai rischi legati alla sicurezza fisica, vanno tenuti in considerazione anche quelli legati alla salute mentale: i lavoratori che devono adeguarsi al ritmo e al livello di lavoro di un *cobot* intelligente potrebbero trovarsi pesantemente sotto pressione e, d'altra parte, la sensibile riduzione del contatto con i colleghi "umani" potrebbe affievolire il necessario confronto e sostegno sociale. Anche da questo punto di vista, è necessaria un'attenzione costante che permetta di intervenire a tutela del benessere e della salute dei lavoratori con azioni mirate, laddove necessario.
Innegabili vantaggi possono essere rinvenuti anche nell'uso di dispositivi di protezione individuale intelligenti, che consentono il monitoraggio in tempo reale dei pericoli e permettono così anche la segnalazione tempestiva di allarmi in caso di problemi di salute e malori, incidenti, esposizioni dannose. Al contrario, il malfunzionamento del sistema o la mancata integrità dei dati raccolti può al contrario essere causa di infortuni o malattie. Inoltre, tutte le informazioni e i dati raccolti con questo metodo potrebbero essere utilizzati dalle organizzazioni al fine di individuare interventi utili per la sicurezza e salute sul lavoro e per migliorare i processi organizzativi e addirittura la produttività. È pur vero, tuttavia, che nel gestire la grande quantità di dati personali - anche particolari o "sensibili" - che potrebbe essere generata, occorrono sistemi e strategie efficaci e, soprattutto, decisioni etiche.
Non va, dunque, sottovalutato il rischio di un controllo a distanza dei lavoratori attraverso l'uso dell'IA. E questo non soltanto per il rispetto della normativa sul punto dettata dallo Statuto dei Lavoratori e dunque dei diritti fondamentali, ma anche per la sicurezza e salute dei lavoratori.
Simili tecnologie possono essere, infatti, in grado di assumere decisioni che riguardano l'organizzazione e lo svolgimento del lavoro, definendo i processi e ottimizzando la produttività, ma di fatto escludendo qualsiasi intervento del lavoratore, che può trovarsi a vedere ridotti i margini di autonomia nell'esecuzione della prestazione e a subire decisioni che hanno un impatto diretto sulla propria vita lavorativa e, dunque, quotidiana (8). È per questo che l'adozione di tecnologie di intelligenza artificiale dovrebbe necessariamente richiedere un coinvolgimento dei lavoratori per il miglioramento e la messa a punto dei sistemi, tanto più che questi si nutrono di dati e informazioni rilasciati, più o meno consapevolmente, dai lavoratori stessi (9).
A ben guardare, infatti, l'art. 4 dello Statuto dei Lavoratori limita il potere di controllo a distanza del datore di lavoro, ma non tocca invece il problema delle decisioni automatizzate e del "nuovo" potere direttivo, che nei casi più estremi può essere esercitato dalla macchina.
La protezione dei lavoratori in questa fase non può che venire, dunque, da una lettura integrata dello Statuto dei Lavoratori e della normativa in materia di trattamento dei dati personali (10). Il

(8) Si pensi, ad esempio, al settore della logistica, laddove un algoritmo può assumere decisioni automatizzate e dettare direttive al lavoratore su tempi e tragitti, comprimendone la sfera di libertà e discrezionalità. Sul punto, si veda Stanzione, *Amazon in Lombardia una battaglia vinta*, Idea diffusa, marzo 2018, 4.

(9) Secondo la Confederazione europea dei sindacati (CES), solo il 23% dei rispondenti ad una ricerca a livello europeo ha segnalato che l'introduzione di nuove tecnologie potenzialmente utilizzabili per monitorare prestazioni e comportamenti o il tema della protezione dei dati personali sono stati oggetto di informazione e consultazione a livello aziendale. CES, *Digitalizzazione e partecipazione dei lavoratori: l'opinione di sindacati, organismi di rappresentanza aziendali e lavoratori delle piattaforme digitali europei*, settembre 2018.

(10) Ci si riferisce ovviamente a Reg. UE 2016/679 e D.Lgs. n. 196/2003. Non potendo in questa sede approfondire il tema, si rimanda a Ingrao, *Il controllo a distanza sui lavoratori e la nuova disciplina privacy: una lettura integrata*, Cacucci, 2018.

Regolamento, infatti, oltre a dettare obblighi in capo al datore di lavoro nella sua qualità di titolare del trattamento, riconosce specifici diritti agli interessati e, dunque, in questo caso ai lavoratori: il diritto di accesso, rettifica, cancellazione, limitazione del trattamento, il diritto di opposizione ed i diritti in tema di trattamento automatizzato costituiscono, allo stato attuale, una protezione importante, anche con riguardo a monitoraggio e controllo a distanza (11).

Intelligenza artificiale, il quadro normativo europeo

Resta necessaria una normativa specifica in materia di intelligenza artificiale, che permetta di governare il fenomeno e non - come purtroppo spesso accade, soprattutto in tema di innovazione e tecnologia - di rincorrerlo.

Il 21 aprile 2021, la Commissione europea ha presentato una proposta di regolamento che stabilisce regole armonizzate in materia di intelligenza artificiale e modifica alcuni atti legislativi dell'Unione. Appare immediatamente chiaro, sin dal primo "Considerando", come lo scopo del regolamento proposto sia "migliorare il funzionamento del mercato interno istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, la commercializzazione e l'uso dell'intelligenza artificiale in conformità ai valori dell'Unione" (12).

Nel medesimo giorno di uscita della proposta di Regolamento, inoltre, la Commissione ha proposto:

— un "Piano coordinato di revisione dell'intelligenza artificiale 2021" (13), che pone le basi affinché la Commissione e gli Stati membri collaborino nell'attuazione di azioni congiunte ed eliminino la frammentazione dei programmi di finanziamento, delle iniziative e delle azioni intraprese a livello dell'UE e dei singoli Stati membri;

— il "Regolamento del Parlamento europeo e del Consiglio relativo alle macchine" (14), che dovrebbe sostituire la Direttiva 2006/42/CE del 17 maggio 2006 relativa alle macchine, che garantisce la libera circolazione delle macchine all'interno del mercato UE ed assicura un alto livello di protezione per gli utenti e altre persone esposte e stabilisce i requisiti di sicurezza dei prodotti.

Tali proposte sono il risultato di un lungo dibattito europeo sulla necessità di costruire un mercato UE dell'IA che sia affidabile, sicuro e rispettoso dei diritti fondamentali (15). Per questa ragione, la proposta di regolamento riguarda ogni aspetto dell'IA nel suo ciclo di vita, dallo sviluppo, all'immissione sul mercato, all'uso e tutti i soggetti coinvolti nello svolgimento di tali attività, quali fornitori, utenti, distributori, importatori o rivenditori. Ciò vale anche per i soggetti situati al di fuori dell'Unione Europea nel caso in cui mettano in servizio sistemi di IA nell'Unione Europea o utilizzino *output* derivanti da sistemi di IA operanti nell'Unione Europea. In tale impostazione risulta evidente il parallelismo con gli effetti extraterritoriali del GDPR (16).

In estrema sintesi, il Regolamento prevede una classificazione dei prodotti che utilizzano completamente o parzialmente software di IA in base al rischio di impatto negativo su diritti fondamentali quali dignità umana, libertà, uguaglianza, democrazia, diritto alla non discriminazione, protezione dei dati e, non ultimo, salute e sicurezza.

L'impostazione normativa risulta, dunque, basata sul rischio: quanto più il prodotto determina dei rischi per i diritti sopra elencati, tanto più severe sono le misure adottate per eliminare o mitigare l'impatto negativo, sino al divieto dei prodotti incompatibili con questi diritti.

Sono, dunque, vietati:

— prodotti suscettibili di distorcere materialmente il comportamento di una persona in un modo che provochi o possa provocare a sé o ad altri un danno fisico o psicologico;

(11) Artt. 15-22, Reg. UE 2016/679, cit.

(12) 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 final 2021/0106 (COD). Il testo integrale in lingua italiana può essere reperito a questo link: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0006.02/DOC_1&format=PDF.

(13) https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review.

(14) Il titolo originale della proposta al momento disponibile solo in inglese è "*Regulation of the European Parliament and of the Council on machinery product*".

(15) Si ricorda che il nuovo approccio europeo fa seguito ad una serie di iniziative intraprese negli ultimi anni, tra cui:

- la consultazione pubblica sul Libro Bianco sull'Intelligenza Artificiale (COM 2020) 65 final del 19 febbraio 2020);
- le Linee guida etiche finali per un'intelligenza artificiale affidabile, del Gruppo ad alto livello sull'intelligenza artificiale, pubblicate l'8 aprile 2019;
- il Rapporto sulla responsabilità per l'Intelligenza Artificiale e altre tecnologie emergenti, del Gruppo di esperti sulla responsabilità e le nuove tecnologie, pubblicato il 21 novembre 2019;
- la Dichiarazione di cooperazione sull'intelligenza artificiale, firmata da 25 paesi europei il 10 aprile 2018, che si basa sui risultati e sugli investimenti della comunità europea della ricerca e delle imprese nell'IA e stabilisce le basi per il Piano coordinato sull'IA.

(16) Reg. UE 2016/679, art. 3.

— prodotti che consentano di valutare o classificare l'affidabilità delle persone fisiche sulla base del loro comportamento sociale o di caratteristiche personali o della personalità note o previste, determinando un trattamento dannoso o sfavorevole.
In linea di massima vietati, i sistemi di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico a fini di attività di contrasto sono ammessi sotto il controllo delle autorità e in casi eccezionali, come la ricerca mirata di vittime di reati, la prevenzione di minacce specifiche o attacchi terroristici (17).
La proposta di regolamento stabilisce, inoltre, una specifica categoria per i sistemi di IA ad alto rischio e cioè quelle tecnologie che presentano un rischio significativo di provocare danni e il cui uso è pertanto consentito solo in presenza di specifici controlli di sicurezza, secondo l'approccio che la normativa europea ha già adottato in materia di sicurezza dei prodotti e gestione del rischio. Tra questi possono rientrare, ad esempio, il *credit scoring*, i sistemi relativi giustizia e sicurezza sociale o a infrastrutture pubbliche essenziali, i sistemi applicati a dispositivi medici e ad altri dispositivi regolamentati o, ancora, al trasporto (18).
Per i sistemi di IA ad alto rischio, la Proposta di Regolamento prevede una serie di controlli tecnici che il fornitore è tenuto ad effettuare, anche per garantire la sicurezza dello stesso per tutto il suo ciclo di vita e la trasparenza nei confronti dell'utente, anche nella redazione di istruzioni e manuali d'uso.
Nel quadro di responsabilizzazione di produttori e fornitori, la proposta prevede anche un obbligo per questi soggetti di verificare e dichiarare la conformità alla normativa prima dell'immissione sul mercato, di istituire, implementare e mantenere un sistema di monitoraggio successivo all'immissione sul mercato e di adottare immediate azioni correttive in caso di non conformità.
Il parallelismo con il GDPR nell'impostazione del Regolamento IA è evidente anche dall'impianto sanzionatorio: sono previste, infatti, sanzioni amministrative pecuniarie fino a 10 o 30 milioni di euro oppure fino al 2% o al 6% del fatturato globale dell'anno finanziario precedente, a seconda della natura e della gravità della violazione (19).
La Proposta, così come il GDPR, utilizza un approccio basato sul rischio e sulla *accountability* e, dunque, sulla responsabilizzazione dei soggetti coinvolti.
Non sono previste, invece, norme specifiche in materia di lavoro e sicurezza (20), per le quali, in assenza di un Regolamento dedicato, dovrà adottarsi una lettura integrata delle diverse discipline coinvolte, relative all'intelligenza artificiale, alla protezione delle persone fisiche circa il trattamento dei dati personali e, ovviamente, alla protezione (nazionale) dei lavoratori.

(17) Art. 5, Proposta di Regolamento, cit.
(18) I sistemi ad alto rischio, di cui all'art. 6 della Proposta di Regolamento, sono elencati all'Allegato III.
(19) Il parallelismo evidenziato è tra art. 71 della Proposta, cit., e Reg. UE 2016/679, art. 83.
(20) La circostanza è stata aspramente criticata dall'ETUI - European Trade Union Institute, cfr. Ponce De Castillo, *The AI Regulation: entering an AI regulatory winter? Why an ad hoc directive on AI in employment is required*, ETUI, 2021.07, https://www.etui.org/publications/ai-regulation-entering-ai-regulatory-winter.