

## INTRODUZIONE

Occorre dunque interrogarsi, in un'ottica non solo reattiva ma anche proattiva, su quale *lex robotica* possa efficacemente governare i nuovi fenomeni, e contemperare le meritevoli istanze di tutela dell'individuo con la pur centrale esigenza di non gravare il settore dell'*high tech* di eccessivi oneri che rischiano di soffocare lo sviluppo tecnologico e scientifico.

Esempio paradigmatico della necessità di un tale bilanciamento può essere rinvenuto nel settore *dell'automotive*, che costituisce forse uno dei più interessanti *case study* per l'analisi delle generali problematiche di mediazione giuridica poste dall'avvento dell'A.I., in relazione sia agli specifici aspetti regolatori, di natura prevalentemente tecnica (requisiti di omologazione, standard di sicurezza, gestione dei dati relativi al traffico, ecc.), sia alle norme in materia di responsabilità per danni da utilizzo, detenzione, proprietà, produzione e commercializzazione di veicoli autonomi.

In tale processo, lo sviluppo di indirizzi interpretativi e/o di nuove eventuali norme non potrà non tenere in considerazione il fatto che, nonostante l'eco mediatica di notizie di (invero rari) incidenti morali imputabili alla circolazione di auto *self-driving* o *driverless*, la diffusione di veicoli autonomi potrà condurre ad una significativa diminuzione del numero dei sinistri.

Tale prospettiva pone, dunque, l'esigenza di evitare di imporre al settore standard e norme di responsabilità eccessivamente penalizzanti, che potrebbero ritardare o disincentivare l'avvento di nuove tecnologie il cui impatto sarà, a livello di macro-fenomeno, comunque benefico. Si pensi alle diffuse tesi di coloro che, richiamando l'abusato "dilemma del carrello", invocano norme e standard che impongano all'auto *driverless* la capacità di valutare il numero di bambini a bordo del

veicolo e su strada, onde determinare quali soggetti siano più sacrificabili in caso di manovre d'emergenza; laddove siffatte imposizioni (oltre ad essere di dubbia liceità) rischierebbero di causare significativi ritardi nella diffusione dei veicoli autonomi.

Generale problema, fra gli altri, attiene a quanto delegare a norme nuove e quanto considerare, invece, risolvibile in sede interpretativa, anche nell'ottica delle dimensioni ormai globali del problema, che dovrebbero imporre soluzioni il più possibile uniformi anche a livello trans-nazionale.

Le innovazioni normative potrebbero, in tal senso, riguardare le previsioni, anche regolamentari ed "amministrative", concernenti gli standard tecnici di settore (criteri di omologazione, regolamentazioni operative, requisiti di commercializzazione, ecc.), idonee a recepire e modulare anche le istanze tecno-etiche reputate meritevoli, limitando, invece, quanto più possibile il ricorso ad innovazioni normative settoriali, ad esempio in materia di responsabilità.

Da un lato, dovremo, allora, riflettere forse sulla necessità di introdurre nuovi doveri di preventiva "omologazione" dei sistemi di A.I., prima che essi possano essere immessi in commercio ed impiegati; tanto, specialmente con riguardo ai settori reputati ad alto rischio (sanità, mobilità, ecc.) o comunque ad elevato impatto sui diritti dell'individuo. Si pensi, a titolo esemplificativo, ai prospettati impieghi dell'A.I. nell'attività anche provvedimentale della pubblica amministrazione e nell'amministrazione della giustizia.

Diversa, ma connessa problematica è quella relativa alla opportunità di congegnare nuove forme di tutela dei "diritti della macchina", se e quando ad una qualche intelligenza non umana sarà accordato lo *status* di persona, o uno *status* comparabile, o quanto meno il riconoscimento di taluni diritti o tutele in quanto "pensante".

Per quanto concerne il mondo del lavoro occorre dire che gli algoritmi, se non opportunamente progettati e monitorati, possono riflettere e amplificare i pregiudizi umani presenti nei dati con cui vengono addestrati. Questo fenomeno è noto come "bias algoritmico" e può portare a discriminazioni e disparità nel processo decisionale automatizzato. I pregiudizi possono essere presenti nei dati storici utilizzati per l'addestramento degli algoritmi, nei criteri di valutazione selezionati o nelle variabili considerate rilevanti per prendere decisioni. Inoltre, la mancanza di flessibilità degli algoritmi può comportare l'esclusione di candidati validi ma con profili diversi da quelli predefiniti. Questo rigido processo decisionale può limitare la diversità e l'inclusione nel processo di selezione del personale.

Quanto detto si collega alla teoria del *nudging*. Il concetto di *nudging* si basa sull'idea di influenzare le scelte delle persone in modo indiretto e non coercitivo, utilizzando strategie che le spingano verso comportamenti desiderati senza limitare la loro libertà di scelta. Questo approccio si inserisce nell'ambito dell'economia comportamentale e della scienza della decisione, che studiano come le persone prendono decisioni in contesti reali, considerando i fattori psicologici e sociali che influenzano tali scelte.

Il *nudging* sfrutta i pregiudizi cognitivi e i modelli di pensiero umani per progettare ambienti decisionali che facilitino scelte più vantaggiose per gli individui e per la società nel suo insieme. In pratica, si tratta di strumenti e interventi progettati per guidare le persone verso comportamenti che migliorino la loro salute, benessere finanziario o felicità, ad esempio attraverso la semplificazione delle opzioni, la messa in evidenza di scelte preferibili o la creazione di incentivi positivi. Per affrontare questi problemi, è fondamentale adottare approcci etici e trasparenti nella progettazione e nell'utilizzo degli algoritmi di selezione del personale. È importante considerare la

diversità, l'equità e l'inclusione come principi guida nella creazione e nell'implementazione di tali strumenti. Inoltre, è consigliabile integrare diversi strumenti di valutazione e non basarsi esclusivamente su test psicoattitudinali per ottenere una visione più completa e accurata dei candidati.

In conclusione, mentre l'intelligenza artificiale offre molteplici vantaggi nel processo di selezione del personale, è essenziale prestare attenzione ai potenziali rischi legati ai pregiudizi algoritmici e alla mancanza di flessibilità per garantire un processo decisionale equo, inclusivo e non discriminatorio.

## CAPITOLO PRIMO

### L'evoluzione storica dell'intelligenza artificiale

#### 1.1. Le linee di sviluppo dell'intelligenza artificiale.

La mancanza di una definizione chiara e universale di intelligenza rende difficile stabilire obiettivi e criteri per valutare i progressi nell'I.A. Inoltre, la complessità e la diversità delle prospettive sull'intelligenza umana complicano ulteriormente il quadro.

L'intelligenza umana è un fenomeno estremamente complesso e multifattoriale, che coinvolge una vasta gamma di capacità cognitive, emotive e sociali. Questo rende difficile definire con precisione cosa significhi essere intelligenti. Allo stesso modo, replicare l'intelligenza umana in un sistema artificiale presenta sfide significative, poiché non esiste una formula definita per l'intelligenza<sup>1</sup>.

Alcune definizioni di intelligenza artificiale si concentrano sull'imitazione delle capacità umane, mentre altre si basano su criteri più ampi, come l'abilità di apprendere, adattarsi e risolvere problemi in modo autonomo. Tuttavia, anche queste definizioni possono essere soggette a interpretazioni diverse e possono non catturare completamente la complessità dell'intelligenza umana.

Questa mancanza di chiarezza concettuale può portare a discussioni e dibattiti prolungati nel campo dell'I.A., con diverse scuole di pensiero che avanzano teorie e approcci contrastanti. Tuttavia, è anche importante riconoscere che questa diversità di prospettive può essere fonte di innovazione e progresso, consentendo a ricercatori e sviluppatori di esplorare una vasta gamma di approcci nell'ambito dell'I.A.

---

<sup>1</sup> G. Saraceni, *Artificial intelligence and mental health*, in *Global network journal*, Vol.5 (2023) Issue 2, 370.

Nel complesso, la ricerca sull'intelligenza artificiale beneficia di una riflessione critica e approfondita sulla natura dell'intelligenza umana e sulla sua relazione con le macchine. Questo processo può aiutare a guidare lo sviluppo di sistemi A.I. più sofisticati e adattabili, che possano integrarsi in modo efficace e etico nella società. L'intelligenza artificiale, da qualsiasi prospettiva la si voglia considerare (e vi sono sicuramente molte prospettive e angolazioni diverse per occuparsene) offre spunti scientifici, tecnologici, epistemologici e filosofici molto importanti, dunque, l'interesse diffuso è comprensibile e giustificato.

Tuttavia, nell'entusiasmo generale, iniziano ad emergere anche varie voci di dissenso e per motivi diametralmente opposti. Da un lato vi sono obiezioni di vecchia data alla possibilità di realizzare effettivamente qualche cosa di artificiale che si possa definire intelligente<sup>2</sup>, obiezioni che hanno alimentato per decenni il dibattito scientifico soprattutto nell'ambito delle scienze cognitive. Di segno opposto invece sono le preoccupazioni di coloro che credono che le macchine, grazie alla continua crescita della loro capacità di accumulare ed elaborare conoscenza, potranno raggiungere un punto di "singolarità tecnologica" oltre il quale l'intelligenza umana verrebbe superata da quella artificiale. Secondo una definizione condivisa da molti l'intelligenza artificiale si occupa di realizzare strumenti (software e hardware) che siano capaci di eseguire compiti normalmente associati all'intelligenza naturale.

Questa almeno era la visione di coloro che coniarono il termine "*artificial intelligence*" in occasione del workshop che si tenne a Dartmouth nel 1956 e che viene considerato l'evento "ufficiale" della nascita della disciplina.

Tuttavia, il sogno di costruire macchine che emulassero pensieri e comportamenti umani ha origini ben più lontane. Il sillogismo aristotelico, come noto,

---

<sup>2</sup> A. Leon Gatys, S. Alexander Ecker, M. Bethge, *A neural algorithm of artistic style*, in *JOVIS*, 2016, Vol.16 n.12.

fu il primo sistema di argomentazione logica che a partire da delle premesse permettesse di dedurre una conclusione, fornendo così un primo esempio di ragionamento formale (nel senso che la correttezza del ragionamento dipende dalla sua struttura, o forma, e non dal significato dei termini impiegati) e dunque un primo sistema di ragionamento meccanizzabile.

Bisogna però aspettare fino al XIII secolo con Ramon Llull (1232-1316) per avere una prima formulazione esplicita dell'idea di calcolo meccanico. Llull con la sua *Ars combinatoria* (secondo la denominazione di Leibniz) non solo propose un linguaggio logico algebrico per rappresentare la conoscenza e generare nuove verità a partire da premesse note, ma fornì anche una implementazione, come diremmo oggi, del suo modello di calcolo.

Llull realizzò infatti una macchina fisica costituita da dischi di carta concentrici che, ruotati opportunamente, permettevano di combinare in vario modo i concetti rappresentati in modo simbolico su di essi mediante lettere. Llull ha avuto una influenza molto importante su René Descartes (1596-1610) e Gottfried Wilhelm Leibniz (1646- 1716), due filosofi che hanno avuto un'enorme importanza nella storia del calcolo. Descartes, probabilmente sollecitato dagli automi meccanici visibili a Parigi e nelle altre città europee, si interessò al rapporto fra corpo e mente e nel *Discorso sul metodo* del 1637 analizzò la possibilità di distinguere un essere umano da un automa che si comportasse come tale<sup>3</sup>.

Questo interesse per gli aspetti cognitivi collegabili agli automi rende Descartes in qualche modo l'antesignano dell'intelligenza artificiale.<sup>4</sup> Quanto a Leibniz, dopo aver dimostrato con la sua calcolatrice<sup>7</sup> che i lavori di calcolo possono

---

<sup>3</sup> R. Descartes, *Discorso sul metodo*, a cura di De Lucia, Roma, 1999.

<sup>4</sup> G. Bara, *Scienza cognitiva. Un approccio evolutivo alla simulazione della mente*, Torino, 1982.

essere utilmente fatti da una macchina invece che dagli esseri umani, si spinge ad immaginare un linguaggio simbolico universale, una “*Characteristica universalis*”, nel quale si possano esprimere tutti i concetti possibili, matematici, scientifici e anche metafisici. Un tale linguaggio fornirebbe lo strumento per la formalizzazione di un qualsiasi problema o controversia filosofica che un “*Calculus raziocinator*” permetterebbe poi di risolvere in modo meccanico mediante il famoso “*calculemus*”, esattamente come con il calcolo si può risolvere un problema matematico.

Il “*Calculus raziocinator*” di Leibniz, dunque, con il suo tentativo di meccanizzare il ragionamento umano in generale, può sicuramente essere visto come un primo tentativo di realizzare una forma di macchina “intelligente”.<sup>5</sup>

Il filo delle macchine intelligenti passa per vari altri personaggi importanti, da Blaise Pascal (1623-1662) a Charles Babbage (1792-1871) e Ada Augusta Byron (1815- 1851), per arrivare ad Alan Mathison Turing (1912-1954), il celebre matematico inglese che oltre ad essere stato, con la macchina che porta il suo nome, uno dei padri fondatori della moderna teoria della computazione - e quindi dei computer per come noi li usiamo oggi - può a buon diritto essere considerato anche uno dei fondatori dell'intelligenza artificiale ante litteram. Turing, infatti, in vari scritti degli anni '40 e '50 del secolo scorso, affronta il tema delle “macchine pensanti” e fornisce dei contributi assolutamente importanti e incredibilmente attuali.

Procedendo in ordine cronologico, in questi anni densissimi di avvenimenti per quanto riguarda la storia dell'informatica e dell'intelligenza artificiale, nel 1947 Turing tiene un seminario pubblico a Londra nel quale, per la prima volta, si parla di “computer intelligence”. E questo viene fatto nel senso, molto moderno, di una macchina che possa imparare dall'esperienza: «*What we want is a machine that can*

---

<sup>5</sup> G. W. Leibniz, *De scientia universalis seu calculo philosophico*, in Jerhardt (a cura di), *Die Philosophischen Schriften von Gottfried Wilhelm Leibniz*, Vol. VII, Berlin, 1890.

*learn from experience ... the possibility of letting the machine alter its own instructions provides the mechanism for this»*<sup>6</sup>. Il contributo più famoso di Turing all'intelligenza artificiale è però il test che porta il suo nome e che in realtà venne da lui chiamato "imitation game".

Già al tempo di Turing in molti iniziavano a porsi la domanda, attuale ancora oggi, relativa alla possibilità che le macchine potessero pensare. Si tratta, come abbiamo visto, di una domanda che è alla base della definizione stessa di intelligenza artificiale, e vi sono molte risposte possibili, derivanti da posizioni filosofiche completamente diverse.

Nel suo famoso articolo *Computing machinery and intelligence* il matematico inglese fornisce una sua risposta indiretta alla domanda, proponendone una versione alternativa.<sup>7</sup> È utile leggere cosa scrive Turing: «*I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think". The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistica! survey such as a Gallup poli. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words. The new forms of the problem can be described in terms of a game which we call the imitation game*».

---

<sup>6</sup> A. Turing, *Lecture to the London Mathematical Society on 20 February 1947*, 1947, 106-124, in Carpenter, Doran, *A.M. Turing's ACE Report of 1946 and Other Papers*, Cambridge, 1986.

<sup>7</sup> A. Turing, *Computing machinery and intelligence*, 1950.

Quindi, argomenta Turing, invece di preoccuparci di rispondere a una domanda per lo meno ambigua, se non mal posta, cerchiamo di capire se una macchina è in grado di superare il test descritto nell'*imitation game*. Questo prevede un interrogatore, chiamiamolo A, che può porre qualsiasi domanda voglia ad un altro essere umano B una macchina C. A riceve le risposte di B e C senza però sapere chi dei suoi due interlocutori è l'essere umano. Se alla fine della sessione di domande A non è in grado, dalle risposte ottenute, di distinguere la macchina dall'essere umano, la macchina ha passato il test, si può definire "intelligente". Nello stesso scritto in cui viene proposto l'*imitation game* Turing analizza e discute varie possibili obiezioni alla sua concezione di macchina intelligente, obiezioni che in larga parte sono poi state effettivamente formulate nei decenni successivi.

Fra di esse forse la più importante riguarda l'assenza della coscienza nella macchina: se la macchina si comporta esternamente in modo intelligente questo non implica che essa effettivamente lo sia. Anche in questo caso la risposta di Turing è notevole (e drammaticamente attuale): facciamo con le macchine quello che facciamo con gli esseri umani, ossia assumiamo *l'educata convenzione* che tutti coloro che dimostrano esternamente una qualche forma di pensiero, pensino davvero.

Negli stessi anni in cui Turing elaborava le sue idee veniva posto un altro mattone fondamentale per la futura disciplina: McCulloch e Pitts nel 1943 proposero il loro modello di neurone artificiale, quello che oggi è alla base delle reti neurali, suggerendo che una rete fatta di tali neuroni potesse apprendere, anche se la formalizzazione del modello di apprendimento fu fatta sei anni più tardi da Hebb con il modello che porta il suo nome e che usiamo ancora oggi.

Le basi, dunque, erano già presenti alla fine degli anni '50 del secolo scorso. Nei decenni successivi si è assistito, con alti e bassi, allo sviluppo scientifico analogo

a quello di tanti altri settori, ovviamente con contributi molto importanti, dai linguaggi di programmazione specifici per l'intelligenza artificiale ai primi sistemi deduttivi basati sulla logica e i primi sistemi esperti (ad esempio MYCIN in ambito medico), fino allo sviluppo delle reti neurali con l'algoritmo di *back propagation* e in tempi più recenti con il *deep learning*.

Tuttavia, non c'è stato, neanche negli ultimissimi anni, un punto di singolarità nella ricerca scientifica che abbia provocato un balzo in avanti delle conoscenze, anzi vi sono stati periodi di grande crisi, incluso il cosiddetto “inverno dell'intelligenza artificiale” degli anni '70 e '80 che, dopo le aspettative forse eccessive degli anni '60, ha visto la delusione di molti per i mancati risultati e la riduzione drastica dei finanziamenti<sup>8</sup>.

Dunque, tornando alla domanda iniziale, perché abbiamo oggi questo enorme interesse, a tutti i livelli, per l'intelligenza artificiale? Cosa è successo negli ultimi anni per arrivare all'esplosione, anche mediatica, di questi temi? La risposta va cercata in una concomitanza di tre fattori che hanno innescato dei meccanismi di rinforzo positivo e creato una sorta di circolo virtuoso per l'efficacia delle applicazioni d'intelligenza artificiale. Il primo fattore importante di cui tener conto è legato ovviamente alla ricerca scientifica e agli sviluppi di alcuni settori, primo fra tutti il machine learning: passare da un modello algoritmico tradizionale, nel quale il programmatore “dice” esattamente alla macchina tutto quello che deve fare, a modelli di calcolo sub-simbolico, nei quali la macchina è in grado di apprendere dall'analisi di opportuni (grandi) insiemi di dati di training, ha permesso di affrontare con successo problemi che con le tecniche tradizionali risultavano difficilmente trattabili. Inoltre, il

---

<sup>8</sup> C. Casonato, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo*, Speciale/2019 pp: 101.