



Classificazione Decimale Dewey:

340.028563 (23.) DIRITTO. INTELLIGENZA ARTIFICIALE

FRANCESCA MORGANTI

**INTELLIGENZA
ARTIFICIALE E TUTELA
DELLA SALUTE
OPPORTUNITÀ E SFIDE
PER IL DIRITTO COSTITUZIONALE**





©

ISBN
979-12-218-0953-4

PRIMA EDIZIONE
ROMA 3 OTTOBRE 2023

*Ai miei nonni,
con affetto e gratitudine*

INDICE

- 9 **Capitolo I**
Intelligenza artificiale e tutela dei diritti fondamentali
1.1. Intelligenza artificiale: le ragioni (e le conseguenze) di una scelta terminologica, 9 – 1.1.1. Sui pericoli presentati dai «pappagalli stocastici», 15 – 1.2. Intelligenza artificiale e diritto: è necessario un «pensiero costituzionale», 17 – 1.2.1. *Privacy e no: i dati e la tenuta del sistema democratico*, 18 – 1.2.2. *Intelligenza artificiale, digitalizzazione dei servizi e principio di eguaglianza*, 21 – 1.2.3. *Sistemi autonomi e robot intelligenti: verso una nuova rivoluzione industriale?*, 24 – 1.2.4. *Soggettività giuridica e personalità elettronica*, 27 – 1.3. «*High performance*», «*flexibility*», «*understandability*»: dai sistemi esperti ai programmi di apprendimento automatico, 29 – 1.3.1. *Perché la protezione dei dati personali non basta*, 33 – 1.3.2. *Alcuni documenti unionali in tema di intelligenza artificiale*, 35 – 1.4. La Proposta di Regolamento «che stabilisce regole armonizzate sull'intelligenza artificiale»: punti di forza e criticità, 38 – 1.4.1. *Approccio basato sul rischio e "duttilità" dei sistemi di intelligenza artificiale*, 45.
- 49 **Capitolo II**
Applicazioni medico-sanitarie dell'intelligenza artificiale e diritto fondamentale alla salute
2.1. Salute e sanità digitale nel Piano nazionale di ripresa e resilienza, 49 – 2.2. Intelligenza artificiale e medicina: un'introduzione..., 51 – 2.2.1. ... *e un esempio operativo*, 55 – 2.3. Discriminazione algoritmica e accesso alle cure in condizioni di eguaglianza, 59 – 2.3.1. *Alcune considerazioni sugli algoritmi di stratificazione, tra esigenze di contenimento della spesa sanitaria e rischi di discriminazione indiretta*, 64 – 2.4. Opacità dei sistemi di intelligenza artificiale e consenso informato, 67 – 2.4.1. *Black-box medicine e responsabilità (cenni conclusivi)*, 69.
- 73 *Bibliografia*

CAPITOLO I

INTELLIGENZA ARTIFICIALE E TUTELA DEI DIRITTI FONDAMENTALI

SOMMARIO: 1.1. Intelligenza artificiale: le ragioni (e le conseguenze) di una scelta terminologica – 1.1.1. *Sui pericoli presentati dai «pappagalli stocastici»* – 1.2. Intelligenza artificiale e diritto: è necessario un «pensiero costituzionale» – 1.2.1. *Privacy e no: i dati e la tenuta del sistema democratico* – 1.2.2. *Intelligenza artificiale, digitalizzazione dei servizi e principio di eguaglianza* – 1.2.3. *Sistemi autonomi e robot intelligenti: verso una nuova rivoluzione industriale?* – 1.2.4. *Soggettività giuridica e personalità elettronica* – 1.3. «High performance», «flexibility», «understandability»: dai sistemi esperti ai programmi di apprendimento automatico – 1.3.1. *Perché la protezione dei dati personali non basta* – 1.3.2. *Alcuni documenti unionali in tema di intelligenza artificiale* – 1.4. La Proposta di Regolamento «che stabilisce regole armonizzate sull'intelligenza artificiale»: punti di forza e criticità – 1.4.1. *Approccio basato sul rischio e «duttilità» dei sistemi di intelligenza artificiale.*

1.1. Intelligenza artificiale: le ragioni (e le conseguenze) di una scelta terminologica

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or

more of these problems if a carefully selected group of scientists work on it together for a summer.⁽¹⁾

La nozione di «intelligenza artificiale» è piuttosto difficile da circoscrivere, e altrettanto complesso, anche per questo, è ripercorrerne la storia. Si potrebbe sostenere, con una semplificazione estrema, che Ada Lovelace ne abbia predetto la possibilità⁽²⁾, Alan Turing ne abbia definito gli scopi⁽³⁾, e John McCarthy⁽⁴⁾ – tra gli organizzatori, assieme a Marvin Minsky⁽⁵⁾, Nathaniel Rochester⁽⁶⁾ e Claude Shannon⁽⁷⁾, del primo *workshop* sull'intelligenza artificiale, tenutosi al Dartmouth College nell'estate del 1956 – abbia posto le basi per il suo sviluppo.

Fu proprio McCarthy, tra l'altro, a coniare l'espressione «intelligenza artificiale», che pure non lo convinceva del tutto: «*I had to call it something, so I called it "Artificial Intelligence", and I had a vague feeling*

(1) J. MCCARTHY, M.L. MINSKY, N. ROCHESTER, C.E. SHANNON, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, 31 agosto 1955.

(2) Nel tradurre e annotare uno scritto di L.F. MENABREA sulla c.d. macchina analitica (*analytical engine*) di Charles Babbage, pubblicato come *Sketch of the Analytical Engine Invented by Charles Babbage, Esq.*, in R. TAYLOR (a cura di), *Scientific Memoirs, Selected from the Transactions of Foreign Academies of Science and Learned Societies, and from Foreign Journals*, Londra, 1843, pp. 666 ss., Ada Lovelace – celandosi dietro le sole iniziali – valorizzò l'attitudine dei numeri, dato il loro carattere simbolico, a rappresentare entità diverse dalla mera quantità, e predisse che macchinari modellati sull'*analytical engine*, considerato come il primo prototipo di *computer* meccanico a finalità generale, sarebbero stati in grado, in futuro, di svolgere compiti assai complessi, e addirittura di «*compose elaborate and scientific pieces of music of any degree of complexity or extent*» (p. 694).

(3) Come sottolinea M.A. BODEN, *L'intelligenza artificiale*, Bologna, 2019, tr. it. di F. Calzavarini, «[d]iversamente da Ada Lovelace, Turing fece propri entrambi gli scopi dell'IA; voleva che le nuove macchine facessero cose utili di cui normalmente si dice che richiedono intelligenza (eventualmente usando tecniche del tutto innaturali), e inoltre che modellassero i processi che avvengono nelle menti a base biologica» (p. 14).

(4) Noto informatico e scienziato cognitivo, John McCarthy (1927-2011) è spesso considerato il «padre fondatore» dell'intelligenza artificiale, della quale, come si dirà nel testo, conio anche il nome. Nel 1955, quando ideò e contribuì a organizzare l'ormai famoso *workshop* sull'*artificial intelligence problem* del Dartmouth College, era un Professore associato di Matematica appena ventottenne.

(5) Marvin Minsky (1927-2016), informatico e scienziato cognitivo, è stato tra i fondatori dell'*Artificial Intelligence Project* – poi *Artificial Intelligence Laboratory* – del Massachusetts Institute of Technology (MIT). È unanimemente considerato come uno dei «padri fondatori» dell'intelligenza artificiale.

(6) Nathaniel Rochester (1919-2001), informatico, è noto per aver progettato – assieme a Jerrier Haddad – l'IBM 701, il primo *computer* commerciale della storia.

(7) Matematico, ingegnere e crittografo, Claude Shannon (1916-2001) è spesso ricordato come il «padre» della teoria dell'informazione.

that I'd heard the phrase before, but in all these years I have never been able to track it down»⁽⁸⁾. Una scelta terminologica (quasi) casuale, dunque, e mirata principalmente a distinguere il nuovo campo «da un'impresa affine, chiamata cibernetica»⁽⁹⁾, ma che ha tuttora ripercussioni sul modo in cui pensiamo e parliamo delle *macchine intelligenti*.

Nella difficoltà di ripercorrere, in particolare entro i confini del presente libello, anche solo i principali tentativi di definire l'intelligenza artificiale⁽¹⁰⁾, può essere utile isolarne due: il primo, il più risalente, è ascrivibile agli stessi “padri” dell'intelligenza artificiale – McCarthy, Minsky, Rochester e Shannon⁽¹¹⁾ – mentre il secondo, ben più recente e “settoriale”, si deve all' *High-Level Expert Group on Artificial Intelligence* (AI HLEG), un gruppo di esperti formato dalla Commissione europea perché – tra l'altro – fosse d'ausilio nell'elaborazione di una strategia eurounitaria per lo sviluppo e la regolazione dell'intelligenza artificiale.

McCarthy e colleghi ipotizzavano che «ogni aspetto dell'apprendimento» o «di qualsiasi altra caratteristica dell'intelligenza» fosse, almeno in linea di principio, «descrivibile tanto precisamente da poter essere

(8) Come riportato dallo stesso McCarthy in una lunga intervista concessa a Nils Nilsson, all'epoca Professore di Informatica a Stanford, il 12 settembre 2007. Una videoregistrazione dell'intervista, dal titolo *Oral History of John McCarthy*, è disponibile in rete (youtu.be/KuU82iz3hi8c).

(9) Come nota M. MITCHELL, *L'intelligenza artificiale. Una guida per esseri umani pensanti*, Torino, 2022, tr. it. di S. Ferraresi, p. 6.

(10) In un recente *report* stilato nell'ambito dello *One Hundred Year Study on Artificial Intelligence* (AI100) della Stanford University, un gruppo di esperti del settore ha nuovamente rilevato «*the lack of a precise, universally accepted definition of AI*», sottolineando però come, controintuitivamente, questa mancanza «*probably has helped the field to grow, blossom, and advance at an ever-accelerating pace*», poiché «*[p]ractitioners, researchers, and developers of AI are instead guided by a rough sense of direction and an imperative to “get on with it”*» (ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE, 2016 Report, disponibile su: ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/ai100report10032016fnl_singles.pdf, p. 12). La definizione operativa accolta nell'ambito dello studio, ad ogni modo, è quella di Nils Nilsson (v. *supra*, nt. 8), per cui «*[a]rtificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment*» (N.J. NILSSON, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge, 2010, p. 13).

(11) Per quanto McCarthy, Minsky, Rochester e Shannon, come detto, siano stati gli organizzatori del primo *workshop* sull'intelligenza artificiale, e siano per questo senz'altro annoverabili tra i “padri” della disciplina, più di frequente come *founding fathers* della medesima sono indicati, accanto ad Alan Turing e agli stessi McCarthy e Minsky, l'informatico Allen Newell (1927-1992), e lo scienziato politico e Premio Nobel per l'economia Herbert Simon (1916-2001), entrambi tra i partecipanti al *workshop*.

simulato da una macchina»⁽¹²⁾. In questo risiedeva l'*artificialità* dell'intelligenza prospettata: nel suo essere non-originale, "riprodotta"; e in questo consisteva «*the artificial intelligence problem*»: nel progetto di realizzare una macchina in grado di comportarsi «*in ways that would be called intelligent if a human were so behaving*»⁽¹³⁾.

Gli interrogativi filosofici più ingombranti (le macchine possono pensare?⁽¹⁴⁾ Qual è la differenza tra «pensare veramente» e «simulare il pensiero»?⁽¹⁵⁾) venivano momentaneamente messi da parte: l'obiettivo era – ed è tuttora, per certi versi – creare macchine *apparentemente* intelligenti, in grado di tenere condotte *apparentemente* intelligenti (leggi: che sarebbero senz'altro considerate tali se tenute da esseri umani, indiscutibilmente – quanto astrattamente – capaci di pensiero).

Oltre mezzo secolo più tardi, questo medesimo approccio si intravede – tra l'altro – nella definizione operativa di «intelligenza artificiale» elaborata dall'AI HLEG, in base alla quale «*[a]rtificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals*»⁽¹⁶⁾. Di nuovo, il *focus* è sui comportamenti *apparen-*

(12) J. MCCARTHY, M.L. MINSKY, N. ROCHESTER, C.E. SHANNON, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, cit., p. 2, tr. dell'A.

(13) Ivi, p. 11.

(14) «*Can machines think?*», com'è noto, è la domanda che si faceva A.M. TURING nel suo *Computing Machinery and Intelligence*, in *Mind*, 1950, pp. 433 ss., dove – a mo' di scherzo, inizialmente – proponeva, per testare l'intelligenza di una macchina, il c.d. *imitation game*, poi conosciuto come *test* di Turing: un "interrogatore" (C) è chiamato a dialogare con un uomo (A), che proverà a ingannarlo, e con una donna (B), che tenterà di aiutarlo; compito dell'interrogatore, che si trova in una stanza separata e riceve le risposte per iscritto, è capire quale dei suoi interlocutori sia di sesso maschile e quale di sesso femminile. «*We now ask the question*», prosegue l'A., «*“What will happen when a machine takes the part of A in this game?” Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, “Can machines think?”*» (p. 434).

(15) Cfr. M. MITCHELL, *L'intelligenza artificiale*, cit., p. 37, dove l'A. riporta una conversazione avuta con sua madre – che le domanda, appunto, «[q]ual è la differenza tra “pensare veramente” e “simulare il pensiero”?» – ispirata da un articolo del *New York Times* sui progressi di Google Translate.

(16) HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A Definition of AI: Main Capabilities and Scientific Disciplines*, 18 dicembre 2018, p. 1. La definizione è riportata, tra l'altro, nella Comunicazione della Commissione su «L'intelligenza artificiale per l'Europa» del 25 aprile 2018, COM(2018) 237 *final*, dove – nella versione italiana, si intende – è tradotta così: «“Intelligenza artificiale” (IA) indica sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi» (p. 1).

temente intelligenti della macchina⁽¹⁷⁾ – *rectius*, del sistema⁽¹⁸⁾ – nonché, e anche questo è rilevante, sul *grado di autonomia* che contraddistingue il sistema medesimo (dove i pressanti interrogativi sull'imputabilità a quest'ultimo, tanto sul piano etico che giuridico, delle condotte che tiene e delle loro conseguenze).

Si è soliti sostenere che l'intelligenza artificiale – il riferimento, in questo caso, è all'ambito di ricerca, non ai singoli sistemi – persegua due obiettivi principali, uno scientifico e l'altro pratico: il primo ha a che vedere con la descrizione e la riproduzione dei meccanismi alla base dell'intelligenza “naturale”, ovviamente al fine di incorporarli in dei calcolatori; l'altro, più prosaicamente, è la creazione di «programmi informatici che eseguano i compiti come, o meglio, degli uomini», perlomeno «senza chiedersi se questi programmi *pensano* davvero come noi»⁽¹⁹⁾.

Eppure, gli ingombranti interrogativi filosofici di cui sopra, in apparenza messi da parte, ciclicamente si inseriscono e reinseriscono nel discorso, altrimenti tecnico-pratico, sull'intelligenza artificiale, e in parte ciò avviene perché si è scelto – più o meno casualmente, come visto – di chiamarla così: una macchina presentata come “intelligente”, come quasi-umana o *super-umana*⁽²⁰⁾, inviterà riflessioni che non sarebbero

(17) *Focus* che, tra l'altro, contraddistingue la stessa definizione di «*artificial intelligence*» utilizzata nell'ambito del *Data Ethics Framework* britannico (una raccolta di principi che dovrebbe “guidare” l'utilizzo dei dati nel settore pubblico): «*AI can be defined as the use of digital technology to create systems capable of performing tasks commonly thought to require intelligence*» (*Data Ethics Framework: Glossary and Methodology*, disponibile su: www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-glossary-and-methodology).

(18) Come sottolineato dall'HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A Definition of AI*, cit., p. 1, «*AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)*»: quando parliamo indiscriminatamente di “macchine”, dunque, ci riferiamo perlomeno – a-tecnicamente – ai sistemi di intelligenza artificiale, siano essi puramente virtuali o incorporati in *hardware*.

(19) M. MITCHELL, *L'intelligenza artificiale*, cit., p. 8.

(20) Secondo Ray Kurzweil, ingegnere capo di Google e teorico della Singolarità («*a future period during which the pace of technological change will be so rapid, its impact so deep, that human life will be irreversibly transformed*»), nel giro di alcuni decenni «*information-based technologies will encompass all human knowledge and proficiency, ultimately including the pattern recognition powers, problem-solving skills, and emotional and moral intelligence of the human brain itself*», ed entro la fine di questo secolo «*the nonbiological portion of our intelligence will be trillions of trillions of times more powerful than unaided human intelligence*». Cfr. R. KURZWEIL, *The Singularity Is Near*, New York, 2005, pp. 7 ss., nonché, per previsioni ancora più “ottimistiche”

stimolate da cenni a “semplici” servizi digitali, dispositivi *hardware* variamente connotati, sistemi anche parzialmente autonomi⁽²¹⁾.

Certo, non si riduce tutto al *nomen*, o comunque al lessico impiegato nel descrivere i sistemi di intelligenza artificiale (nel presentarli, appunto, come *intelligenti*, *autonomi*, come *pensanti* e talvolta *parlanti*): sono i compiti stessi assegnati a tali sistemi e da loro svolti – riconoscere immagini, tradurre testi, rispondere a domande, finanche giocare a scacchi⁽²²⁾ o comporre brani musicali⁽²³⁾ – a sembrare, secondo una gradazione crescente, inerentemente umani.

– Kurzweil ci tiene a presentarle come «*neither utopian nor dystopian*» – i suoi scritti e le sue interviste più recenti.

(21) Come nota A. D'ALOIA, *Ripensare il diritto nel tempo dell'intelligenza artificiale*, in A. PAJNO, F. DONATI, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, I, *Diritti fondamentali, dati personali e regolazione*, Bologna, 2022, p. 81, «già solo parlare di intelligenza artificiale, accostare le due parole, crea una sensazione immediata di superamento del confine della rilevanza per l'etica e per il diritto».

(22) Il riferimento, com'è chiaro, è a *Deep Blue*, il *supercomputer* dell'IBM che nel 1997 sconfisse Gary Kasparov, l'allora campione del mondo di scacchi, in una sfida in sei partite.

(23) Nel suo celeberrimo *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante*, Douglas Hofstadter, noto studioso di fisica, scienze cognitive, filosofia della mente e letterature comparate, nonché esperto di intelligenza artificiale, presentava, in coda a oltre settecento pagine di riflessioni assai complesse e “multidisciplinari”, «dieci domande sull'IA e i relativi tentativi di risposta». La prima domanda, la più importante, era: «Un programma di calcolatore sarà mai in grado di scrivere buona musica?», alla quale l'A. rispondeva: «Sì, ma non tanto presto. La musica è un linguaggio di emozioni, e finché i programmi non avranno emozioni complesse come le nostre non sarà possibile che un programma scriva qualcosa di bello. Potranno esservi “contraffazioni”, superficiali imitazioni della sintassi musicale preesistente; ma [...] vi è molto di più nell'espressione musicale di quanto si possa catturare in regole sintattiche. [...] Considero grottesca e un'inaccettabile sottovalutazione della profondità dello spirito umano la posizione di chi pensa, e c'è chi lo pensa, che tra breve potremmo essere in grado di ordinare a una “scatola per musica” preprogrammata, prodotta in serie per una diffusione di massa a basso costo, di tirar fuori dai suoi sterili circuiti pezzi che Chopin e Bach avrebbero potuto scrivere se fossero vissuti più a lungo» (D.R. HOFSTADTER, *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante*, Milano, 1984, ed. it. curata da G. Trautteur e tradotta a più mani, pp. 731 s.). Come raccontato dallo stesso Hofstadter nel corso di un *meeting* svoltosi nel maggio del 2014 a Mountain View, in California, presso il quartier generale di Google, la sua completa fiducia in quella previsione fu scossa già negli anni Novanta, quando il musicista David Cope “scrise” un programma – *Experiments in Musical Intelligence*, o Emi – in grado di creare brani nello stile di Chopin e Bach. Nelle parole di Hofstadter, riportate da M. MITCHELL, *L'intelligenza artificiale*, cit., pp. XIII-XIV: «Mi sedetti al pianoforte e suonai una delle mazurke di Emi “nello stile di Chopin”. Non sembrava proprio Chopin, ma era abbastanza simile a Chopin, e sembrava una musica coerente, tale da turbarmi *profondamente*. [...] Si ha come la sensazione che *nulla* vi sia di più umano al mondo dell'espressione della musica. Nulla! L'idea che la manipolazione di schemi più superficiali possa generare qualcosa, come se provenisse dal cuore di un essere umano, è molto, ma molto inquietante. Tutto questo mi aveva letteralmente sconvolto».

I.I.I. *Sui pericoli presentati dai «pappagalli stocastici»*

In un fondamentale *paper* del 2021, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, Emily M. Bender, Timnit Gebru *et al.* hanno paragonato i modelli linguistici di grandi dimensioni (*large language models*, o LLMs) ad altrettanto ingombranti «*stochastic parrots*», evidenziando la natura *solo apparentemente comunicativa* dei testi che generano e dei dialoghi cui sembrano partecipare⁽²⁴⁾.

I modelli linguistici – sui quali si fonda, tra le altre, l’applicazione *ChatGPT* – sono sistemi riconducibili nel macro-insieme dell’intelligenza artificiale basata sull’apprendimento automatico (*machine learning*). Della differenza tra intelligenza artificiale simbolica, ovvero basata sui modelli, e sub-simbolica, com’è quella del *machine learning*, si dirà più avanti⁽²⁵⁾; quel che preme sottolineare, in questa sede, è che i *large language models* si servono di reti neurali – così chiamate, intuitivamente, perché consistono in insiemi di *neuroni simulati* variamente interconnessi – per “apprendere” da grandi quantità di dati e svolgere, all’esito dell’addestramento, compiti tradizionalmente e inerentemente umani: traducono e riassumono testi, stilano *report* a partire da appunti, redigono *e-mail*, rispondono a domande, *generano* nuovi scritti, immagini, contenuti audio e video, sulla scorta di “suggerimenti” (*prompts*)⁽²⁶⁾.

Per quanto il “linguaggio” di simili modelli appaia coerente al suo interno e tendenzialmente *on-topic*, avvertono Bender e Gebru, coerenza e intellegibilità sono solo «negli occhi di chi guarda»: come esseri umani, siamo soliti valutare ogni insieme di parole, scritte o pronunciate che siano, in relazione all’intento comunicativo di chi scrive o di chi parla. Prendiamo in considerazione l’esplicito e l’implicito, le espressioni non-verbali tanto quanto quelle verbali; se non conosciamo colui o colei che sta parlando, costruiamo un modello dell’interlocutore e lo utilizziamo per interpretare il suo messaggio.

(24) E.M. BENDER, T. GEBRU *et al.*, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, in *FACCT ’21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York, 2021, pp. 610 ss.

(25) V. *infra*, § 1.3.

(26) Come efficacemente riassunto nell’*explainer* dell’Ada Lovelace Institute relativo ai cc.dd. modelli di base (*foundation models*), disponibile su: www.adalovelaceinstitute.org/resource/foundation-models-explainer/.

Il testo generato da un sistema di intelligenza artificiale non è veicolo di alcun intento comunicativo, non è radicato in alcuna *Weltanschauung*: non ha significato, se non quello attribuitogli *unilateralmente* dagli esseri umani che lo ricevono e interpretano (tendenzialmente costruendo, e questo è un primo problema, un modello antropomorfo e necessariamente fallace dell'interlocutore non-umano)⁽²⁷⁾.

Contrariamente a quel che sembra quando se ne ricevono gli *output*, il sistema si limita a creare stringhe di parole – di simboli – sulla base di calcoli probabilistici: addestrato su *corpora* sterminati, è in grado di valutare la probabilità che a un *token* (carattere, parola...) ne segua un altro – o un altro ancora – analizzando il contesto precedente e, nei modelli bidirezionali e mascherati, quello circostante⁽²⁸⁾. La macchina non si cura del significato dei termini, né li tratta realmente come tali: mette assieme simboli e suoni per come li ha incontrati e “sentiti”, ed è questo che fa di lei uno «*stochastic parrot*».

Se nessuno, tuttavia, tratterebbe come “linguaggio” i garriti di un pappagallo, i contenuti generati dai *large language models* vengono interpretati, elevati a riferimento, impiegati in vario modo. La «fluenza» apparente dei modelli linguistici di grandi dimensioni pone numerosi e concreti rischi, proprio perché «*humans are prepared to interpret strings belonging to languages they speak as meaningful and corresponding to the communicative intent of some individual or group of individuals who have accountability for what is said*»⁽²⁹⁾.

I *corpora* su cui il modello linguistico è addestrato – testi, immagini, tracce: dati – sono sterminati e scarsamente curati, perlopiù raccolti sulla rete. Sono portatori di visioni del mondo egemoniche⁽³⁰⁾, ma po-

(27) E.M. BENDER, T. GEBRU *et al.*, *On the Dangers of Stochastic Parrots*, cit., p. 616. Sul punto v. anche E.M. BENDER, A. KOLLER, *Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data*, in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Stroudsburg (PA), 2020, pp. 5185 ss., dove si sottolinea la distanza intercorrente tra *form* (significante) e *meaning* (a metà, nel caso di specie, tra significato in senso stretto e intento comunicativo), evidenziando come «*the language modeling task, because it only uses form as training data, cannot in principle lead to learning of meaning*», e come sia azzardato, dunque, parlare di “comprensione” o di “sapere” delle macchine.

(28) E.M. BENDER, T. GEBRU *et al.*, *On the Dangers of Stochastic Parrots*, cit., p. 611.

(29) Ivi, p. 617.

(30) Come sottolineato ivi, p. 614, «*[i]n accepting large amounts of web text as “representative” of “all” of humanity we risk perpetuating dominant viewpoints, increasing power imbalances, and further reifying inequality*».

tenzialmente “problematiche”, e disseminati di rappresentazioni e associazioni stereotipiche⁽³¹⁾, oltre che di informazioni non verificate. Le distorsioni dei *set* di dati impiegati in sede di *training* inevitabilmente si ripercuotono sull’operato del sistema e sugli *output* che questo produce (classificazioni, predizioni...). Nelle parole di Birhane e Prabhu, parafrasando Ruha Benjamin⁽³²⁾: «*Feeding AI systems on the world’s beauty, ugliness, and cruelty, but expecting it to reflect only the beauty is a fantasy*»⁽³³⁾.

1.2. Intelligenza artificiale e diritto: è necessario un «pensiero costituzionale»⁽³⁴⁾

C’è un problema di qualità dei dati, ma non solo.

Nella prima bozza delle *Ethics Guidelines for Trustworthy AI* dell’*High-Level Expert Group on Artificial Intelligence*, resa pubblica nel dicembre del 2018, si leggeva, già nelle primissime righe, che l’intelligenza artificiale è «*one of the most transformative forces of our time, and is bound to alter the fabric of society*»⁽³⁵⁾. Per quanto l’osservazione in questione non sia stata inclusa nella versione definitiva del documento⁽³⁶⁾ – *rectius*, sia stata relegata, leggermente modificata e riferita ai soli sistemi cc.dd. sociali, nella nt.

(31) Cfr., *ex multis*, C. BASTA, M.R. COSTA-JUSSÀ, N. CASAS, *Evaluating the Underlying Gender Bias in Contextualized Word Embeddings*, in *Proceedings of the 1st Workshop on Gender Bias in Natural Language Processing*, Stroudsburg (PA), 2019, pp. 33 ss.; K. KURITA *et al.*, *Measuring Bias in Contextualized Word Representations*, *ivi*, pp. 166 ss.; E. SHENG *et al.*, *The Woman Worked as a Babysitter: On Biases in Language Generation*, in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, Stroudsburg (PA), 2019, pp. 3407 ss.; nonché, con riferimento all’impiego di modelli linguistici in ambito medico-sanitario, H. ZHANG *et al.*, *Hurtful Words: Quantifying Biases in Clinical Contextual Word Embeddings*, in *CHIL ’20: Proceedings of the ACM Conference on Health, Inference, and Learning*, New York, 2020, pp. 110 ss.

(32) R. BENJAMIN, *Race after Technology*, Cambridge, 2019, *passim*.

(33) A. BIRHANE, V.U. PRABHU, *Large Image Datasets: A Pyrrhic Win for Computer Vision?*, in *Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision*, Piscataway (NJ), 2021, p. 1541.

(34) Così – condivisibilmente – A. D’ALIOIA, *Ripensare il diritto nel tempo dell’intelligenza artificiale*, *cit.*, p. 93.

(35) HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *Draft: Ethics Guidelines for Trustworthy AI. Working Document for Stakeholders’ Consultation*, 18 dicembre 2018, p. i.

(36) HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *Ethics Guidelines for Trustworthy AI*, 8 aprile 2019.

47⁽³⁷⁾ – essa rimane presupposto dell'intero discorso sull'intelligenza artificiale, e dello stesso sforzo, a livello sovranazionale e non solo, di comprenderla e regolarla (il che vuol dire, tra l'altro, limitarla).

Se è vero, come è vero, che l'impiego su larga scala di sistemi di intelligenza artificiale inevitabilmente modificherà «il tessuto della nostra vita sociale»⁽³⁸⁾ – e in parte lo sta già modificando – è indiscutibile che tali sistemi si collochino ben oltre la soglia di rilevanza per l'etica e, quel che più rileva in questa sede, per il diritto. «La *AI revolution*», è stato scritto, «ha bisogno di essere accompagnata e “sorretta” da un pensiero “costituzionale”, deve produrre una risposta in termini di concettualizzazione di diritti e principi», così come «la rivoluzione industriale ha prodotto l'evoluzione welfarista degli Stati liberali nel XIX secolo e il costituzionalismo sociale del XX secolo»⁽³⁹⁾.

Come è stato per la rivoluzione industriale, la c.d. *AI revolution* è destinata a impattare – in vario modo, e con diversi gradi di intensità – sulla maggior parte delle situazioni oggi protette dalle Costituzioni (e non solo⁽⁴⁰⁾). Starà al diritto pubblico, e al diritto in genere, adattarsi al mutato contesto: sarà necessario ripensare alcune categorie ed elaborarne di nuove, “aggiustare” equilibri rivelatisi fragili, tentare di tenere l'uomo al centro – ammesso che l'antropocentrismo seguiti a essere considerato un valore – accompagnando e *indirizzando* il progresso tecnologico.

Tra le molte questioni che si pongono, ad ogni modo, alcune appaiono più urgenti: vi sono settori della vita sociale (ed economica) nei quali l'impatto dell'intelligenza artificiale è già notevole, e vi sono diritti – e interessi – che sembrano più direttamente minacciati dal progresso recente. Proveremo a individuarne alcuni.

1.2.1. Privacy e no: i dati e la tenuta del sistema democratico

Il discorso sui dati viene incentrato, perlopiù, attorno all'istituto del consenso: quest'ultimo, «espresso mediante un atto positivo [...] con

(37) Dove si legge che «*those systems* [i sistemi sociali, appunto] *have the potential to change our socio-cultural practices and the fabric of our social life*».

(38) V. nt. precedente.

(39) A. D'ALOIA, *Ripensare il diritto nel tempo dell'intelligenza artificiale*, cit., p. 93.

(40) Il riferimento è, tra l'altro, agli strumenti para-costituzionali del diritto internazionale e sovranazionale.

il quale l'interessato manifesta l'intenzione *libera, specifica, informata e inequivocabile* di accettare il trattamento dei dati personali che lo riguardano»⁽⁴¹⁾, rappresenta la principale condizione di liceità del trattamento stesso⁽⁴²⁾. Ma non tutti i dati sono personali⁽⁴³⁾, e non sempre il consenso muove da un'intenzione realmente *libera, specifica, informata e inequivocabile*.

Se l'accesso a un servizio digitale è subordinato alla fornitura (e al consenso al trattamento) di dati personali dell'utente, il servizio *non può considerarsi gratuito*: i dati sono moneta⁽⁴⁴⁾. Quando si tratti di un

(41) Così il Considerando (32) del Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, «relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati)», del 27 aprile 2016, noto come GDPR (*General Data Protection Regulation*). L'enfasi è aggiunta.

(42) Cfr. art. 6, par. 1, lett. a), GDPR («Liceità del trattamento»)

(43) *Ex art.* 4, par. 1, n. 1), GDPR, si qualifica come dato «personale» solo quello «riguardante una persona fisica identificata o identificabile (“interessato”); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo *online* o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale».

(44) *Rectius*: sono concepibili come utilità contrattuali. Nelle parole di Chris Palmer, all'epoca *technology director* della *Electronic Frontier Foundation*: «*You're getting a free service, and the cost is information about you. And Google and Facebook translate that pretty directly into money*» (cfr. E. PARISER, *The Filter Bubble. What the Internet Is Hiding from You*, Londra, 2011, p. 6). Si v., sul punto, il provv. n. 27432 dell'Autorità garante della concorrenza e del mercato, adottato il 29 novembre 2018, che bollava come pratica commerciale «ingannevole» – ai sensi degli artt. 21 ss. del d.lgs. 6 settembre 2005, n. 206 (c.d. Codice del consumo) – il *claim* di gratuità presente all'epoca nella *homepage* di Facebook («Iscriviti – È *gratis* e lo sarà sempre»): «[p]remesso che FB fonda la propria attività economica sulla pubblicità mirata *online* basata sulla profilazione degli utenti», si legge al pt. 55, «la pratica [descritta *supra*] risulta ingannevole in quanto nella schermata di registrazione a FB [...] il Professionista omette informazioni rilevanti di cui il consumatore necessita [...]. Nello specifico, Facebook non informa l'utente con chiarezza e immediatezza in merito alla raccolta e all'utilizzo, a fini remunerativi, dei dati dell'utente da parte del Professionista e, conseguentemente, dell'intento commerciale perseguito, volto alla monetizzazione dei medesimi». La ricostruzione in questione è stata confermata e «ampliata» dal TAR Lazio, sez. I, sent. 10 gennaio 2020, n. 260, dove si sottolineava che «[i]l fenomeno della “*patrimonializzazione*” del dato personale, tipico delle nuove economie dei mercati digitali, impone agli operatori di rispettare, nelle relative transazioni commerciali, quegli obblighi di chiarezza, completezza e non ingannevolezza delle informazioni previsti dalla legislazione a protezione del consumatore, che deve essere reso edotto dello *scambio di prestazioni* che è sotteso alla adesione ad un contratto per la fruizione di un servizio, quale è quello di utilizzo di un *social network*» (pt. 6 della parte in diritto, enfasi aggiunte), e poi dal Cons. St., sez. VI, sent. 29 marzo 2021, n. 2631. Sul punto v., anche per ulteriori riferimenti bibliografici, C. SOLINAS, *Circolazione dei dati personali, onerosità del contratto e pratiche commerciali*

servizio *percepito come essenziale* nell'odierna società dell'informazione – si pensi a «modelli di *business* e servizi nuovi e innovativi, quali le reti sociali (cosiddetti *social network*) e le piattaforme *online* che consentono ai consumatori di concludere contratti a distanza con operatori commerciali», che «la maggior parte dei cittadini dell'Unione utilizza [...] quotidianamente»⁽⁴⁵⁾ – l'intera transazione (dati per accesso) assume contorni di necessità, e la libertà del consenso è più postulata che reale. Non solo: sono ben pochi gli utenti che realmente leggono le pagine e pagine di “termini e condizioni” del servizio; i più si limitano a prestare un consenso «consapevolmente disinformato»⁽⁴⁶⁾.

Chi raccoglie e controlla questi dati «ha in mano formidabili strumenti di potere e di influenza: economica, sociale, politica»⁽⁴⁷⁾. Perché i dati, come detto, hanno un valore economico, e perché – mediante sistemi di intelligenza artificiale⁽⁴⁸⁾ – possono essere (ceduti a terzi e) utilizzati per attività di profilazione e *microtargeting*: le informazioni che l'utente *volontariamente* condivide – nonché, più in generale, tutte le “tracce” che lascia *online* – sono impiegate per «analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti» della persona fisica⁽⁴⁹⁾, così da realizzarne un profilo e indirizzarle, oltre ad annunci pubblicitari mirati, messaggi politici altrettanto “su misura”.

Non solo *fake news*, dunque: anche fenomeni come quello delle bolle di filtraggio (*filter bubbles*)⁽⁵⁰⁾ o delle camere di risonanza (*echo*

scorrette, in *Giur. it.*, 2/2021, pp. 321 ss., in nota alla prima sentenza amministrativa; nonché, più in generale, i contributi raccolti in V. RICCIUTO, C. SOLINAS (a cura di), *Forniture di servizi digitali e «pagamento» con la prestazione dei dati personali. Un discusso profilo dell'economia digitale*, Milano, 2022.

(45) Così il Considerando (1) del Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio, «relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali)», del 19 ottobre 2022, noto come *Digital Services Act*.

(46) Secondo una felice espressione – sulla quale si tornerà *infra* – di C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Dir. pubbl. comp. eur.*, spec./2019, p. 107.

(47) A. D'ALOIA, *Ripensare il diritto nel tempo dell'intelligenza artificiale*, cit., p. 95.

(48) Cfr. M. FASAN, *Intelligenza artificiale e pluralismo: uso delle tecniche di profilazione nello spazio pubblico democratico*, in *BioLaw Journal - Rivista di BioDiritto*, 1/2019, pp. 107 ss.

(49) Secondo la definizione di «profilazione» contenuta nell'art. 4, par. 1, n. 4), GDPR.

(50) L'espressione «*filter bubble*» è stata coniata dall'attivista E. PARISER, *The Filter Bubble*, cit., che la spiega così: «*The new generation of Internet filters look at the things you seem to like – the actual*