

## Costituzione e intelligenza artificiale: un'agenda per il prossimo futuro

Carlo Casonato\*

CONSTITUTION AND ARTIFICIAL INTELLIGENCE: AN AGENDA FOR THE NEAR FUTURE

ABSTRACT: The Italian Constitution was written in 1946 and 1947, when the Turing test was not even developed. Yet, a number of principles contained in it can be effectively applied in order to obtain a human-centered Artificial Intelligence. This article deals with a few constitutional principles (democracy, sovereignty, equality, etc.) in order to identify new rights (such as the right to know about the rationale behind the AI decision, the right to know about the artificial or human nature of the entity I am talking with, the right to a decision not based only on automated processing) that can make AI a valuable resource – and not a threat – for individuals and the community.

KEYWORDS: Artificial Intelligence; Constitutional law; fundamental rights; human-centered AI

SOMMARIO: 1. Oggetto del lavoro: una AI costituzionalmente orientata? – 2. AI: chiarimenti e ambiti di applicazione – 3. AI, democrazia e sovranità popolare (art. 1) – 4. Il lavoro: potenzialità e rischi (artt. 1 e 4) – 5. AI e eguaglianza (art. 3): salute (art. 32) e giustizia (art. 111) – 6. AI e privacy: il rischio per antonomasia – 7. La decisione automatizzata – 7.1 Sostituzione v. assistenza 8. I diritti per una AI costituzionalmente orientata – 9. Conclusioni: le potenzialità della Costituzione italiana nell'orientare una disciplina della AI.

### 1. Oggetto del lavoro: una AI costituzionalmente orientata?

**N**ulla, al momento della scrittura della Costituzione repubblicana, poteva far intuire gli sviluppi scientifici e tecnologici che hanno portato all'intelligenza artificiale (d'ora in poi AI). Tuttavia, come peraltro già accaduto in tanti altri ambiti, la capacità adattiva dei principi contenuti nella Carta ne permette una efficace applicazione anche nei confronti di tale innovativo settore.

In questo contributo, dopo aver definito – per quanto possibile – cosa possa intendersi per AI, passerò in breve rassegna una molteplicità di disposizioni costituzionali per scoprirne le potenzialità tanto in termini di promozione della tecnologia (ex art. 9 Cost.), quanto in termini di protezione da un suo utilizzo sconsiderato. Fra le posizioni catastrofiste e quelle entusiaste, la linea di questo breve scritto privilegerà una via intermedia, tesa a all'individuazione dei principi, delle categorie e degli istituti in grado di dare un orientamento costituzionale alla AI.

---

\* *Professore ordinario di diritto costituzionale comparato, Facoltà di Giurisprudenza, Università degli Studi di Trento e membro del Comitato Nazionale per la Bioetica. Mail: [carlo.casonato@unitn.it](mailto:carlo.casonato@unitn.it). Contributo sottoposto a referaggio.*

## 2. AI: chiarimenti e ambiti di applicazione

L'esistenza di intelligenze non naturali, in senso lato definibili artificiali, è un dato che si ritrova in credenze e riflessioni sparse all'interno di una molteplicità di varianti spazio-temporali, in diversi miti, confessioni religiose, manifestazioni artistiche e speculazioni filosofiche<sup>1</sup>. Il termine AI, tuttavia, è divenuto oggi frequente in riferimento alle capacità di apprendimento, ragionamento e autocorrezione da parte di sistemi informatici in grado di simulare specifiche funzioni dell'intelligenza umana.

La metà degli anni '50 del secolo scorso segna convenzionalmente la nascita della moderna idea di AI<sup>2</sup>. Con il passare del tempo, e dopo il cosiddetto inverno della AI, tale tecnologia ha avuto uno sviluppo imponente e diffuso in una molteplicità di ambiti di applicazione. Dalla comunicazione alla difesa, dai trasporti alla giustizia, dalla medicina alla finanza, dall'insegnamento allo svago non paiono esserci ambiti dell'attività umana svolta con l'assistenza della tecnologia che non siano già stati (e ancor più saranno) pesantemente influenzati dall'utilizzo della AI.

Tale multiforme utilizzo ha condotto all'individuazione di definizioni assai diversificate, tanto che alcuni autori hanno calcolato non meno di un centinaio di definizioni non coincidenti di AI<sup>3</sup>. Così, si è parlato di sistemi informatici puramente *software-based*, come i meccanismi di assistenza vocale, riconoscimento facciale, analisi d'immagine e motori di ricerca che agiscono esclusivamente in una dimensione virtuale, contrapposti a quelli incorporati in dispositivi *hardware*, agenti concretamente nel mondo reale nelle vesti di robot avanzati, autoveicoli autonomi, droni o internet delle cose (*Internet of things: IoT*)<sup>4</sup>. Ancora, solo per fare un secondo esempio, si parla di *weak AI*, con riferimento agli algoritmi in grado di svolgere in maniera molto accurata uno specifico compito, e di *strong AI*, in relazione ai sistemi che, secondo alcuni, saranno in grado, in futuro, di compiere qualunque attività umana come e meglio degli umani. D'altro canto, si è proposto di sostituire il riferimento all'intelligenza con quello alla razionalità, intesa come la capacità di scegliere l'azione migliore da intraprendere per raggiungere un determinato obiettivo, dati determinati criteri da ottimizzare e una certa quantità di risorse disponibili<sup>5</sup>. Nonostante questa varietà, e con un certo grado di approssimazione, una definizione di AI è stata data, a livello europeo, in riferimento ai sistemi che mostrano un comportamento intelligente attraverso l'analisi dell'ambiente in cui sono inseriti e l'adozione di azioni rivolte al raggiungimento di specifici

<sup>1</sup> Nell'Iliade, ad esempio, si narra che Efesto, dio del fuoco, avesse costruito robot che lo potessero assistere negli spostamenti e nelle faccende domestiche. Lo stesso Efesto, avrebbe inoltre forgiato *Telos*, un gigante di rame che Zeus avrebbe regalato a Minosse per difendere Creta dai nemici. Una leggenda cinese narra di Huang Ti, il figlio del cielo, che sarebbe giunto sulla terra attorno al 2500 a.C. con servitori robot. Fra gli altri, M. NEGROTTI, *Towards a general theory of the artificial*, in *AI & Society*, 14, 2000, 268. In generale, S. FRANKLIN, *Artificial Intelligence, General*, in A.L.C. RUNEHOF, L. OVIEDO (eds), *Encyclopedia of Sciences and Religions*, 2013, Dordrecht.

<sup>2</sup> Nel 1955, si tenne una *Session on Learning Machines* in collegamento con una *Western Joint Computer Conference* a Los Angeles; nel 1956, i ricercatori di un *Summer Research Project on Artificial Intelligence* si riunirono presso il Dartmouth College; nel 1958, un simposio sulla *Mechanization of Thought Processes* fu convocato dal *National Physical Laboratory* britannico: N.J. NILSSON, *The quest for Artificial Intelligence*, Cambridge, 2009, 73.

<sup>3</sup> In generale, J. TURNER, *Robot Rules*, Cham, 2019.

<sup>4</sup> Come per le precisazioni seguenti, cfr. il lavoro dell'High-Level Expert Group on Artificial Intelligence nominato dalla Commissione europea su *A definition of AI: Main Capabilities and Disciplines*, Brussels, aprile 2019.

<sup>5</sup> Il profilo definitorio è indagato da M.C. CARROZZA ET AL., *AI: profili tecnologici. Automazione e Autonomia: dalla definizione alle possibili applicazioni dell'Intelligenza Artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, 3, 2019, 243 ss.

obiettivi e connotate da un certo grado di autonomia<sup>6</sup>. Nonostante la sua genericità, questa definizione è sufficientemente esplicativa allo scopo di individuare i profili giuridici idonei a orientare un utilizzo costituzionalmente orientato della AI.

### 3. AI, democrazia e sovranità popolare (art. 1)

La qualità democratica della Repubblica e la appartenenza al popolo della sovranità (ai sensi dell'art. 1 Cost.) sono prime caratteristiche che possono dare indicazioni utili per un impiego della AI che sia centrato sui valori costituzionali<sup>7</sup>.

Il discorso può farsi partire dagli studi secondo cui, già nel 2020, il volume dei dati trattati a livello planetario sarà di 44 zettabyte (unità di misura corrispondente a  $10^{21}$  byte) equivalente, in termini minimamente più comprensibili, a un multiplo di quasi 8 miliardi di volte tutte le informazioni conservate nella biblioteca del Congresso di Washington<sup>8</sup>. Tale impressionante quantità di dati, riferibili alle attività più disparate quotidianamente svolte da ognuno di noi, costituisce una fonte di materia prima preziosissima (l'equivalente del petrolio del futuro, si è detto), la cui gestione però sfugge, per complessità, a qualunque mente umana. Al fine di stratificare i dati e di estrarne, quindi, un qualche significato compiuto (operazione di *data mining*), è necessario l'impiego di sofisticati sistemi di AI, gli unici in grado di compiere il passaggio dai dati all'informazione. Grazie a tali sistemi, alcune (poche) potentissime società sono in grado di svolgere un'operazione di profilazione (*profiling*) che permette di penetrare ad estrema profondità, anche se non sempre con equivalente accuratezza, l'identità personale, individuando i gusti e gli orientamenti complessivi di ognuno di noi<sup>9</sup>. E come accade per la pubblicità commerciale via internet, in cui vengono segnalate offerte sempre più corrispondenti al gradimento individuale al fine di far aumentare i consumi, i sistemi più avanzati di AI sono addestrati a dirigere,

<sup>6</sup> Sono questi i termini della definizione di partenza dell'High-Level Expert Group on Artificial Intelligence *A definition of AI: Main Capabilities and Disciplines*, cit., che, a pag. 6, propone una definizione più articolata: «Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)».

<sup>7</sup> Come in altre parti dello scritto, riprenderò in forma sintetica alcune considerazioni che ho più diffusamente argomentato in *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto Pubblico Comparato ed Europeo*, numero speciale, 2019, 101-130.

<sup>8</sup> Si veda, l'executive summary *Data Growth, Business Opportunities, and the IT Imperatives*, de *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the IoT* (aprile 2014): <https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm> (tutti i siti citati sono stati consultati, da ultimo, il 28 giugno 2019).

<sup>9</sup> Si vedano, ad esempio, le interviste a diversi ricercatori pubblicate periodicamente su *AI Matters, AI profiles: an interview with*. In generale, L. ANTONINI, *Globalizzazione e nuove sfide del costituzionalismo*, in *Diritto Pubblico*, 2, 2019, 319 ss.; P. COSTANZO, *La democrazia digitale (precauzioni per l'uso)*, in *Diritto pubblico*, 1, 2019, 71 ss.

soprattutto attraverso i *social*, un flusso di informazioni di natura politica del tutto omogeneo e corrispondente alle preferenze manifestate dalla singola persona. I potenziali futuri elettori sono così esposti a notizie e segnalazioni, talvolta vere ma sempre più spesso false e costruite ad arte dagli stessi sistemi (dis)informatici, molto limitate, del tutto parziali e ritagliate su misura sulla base delle rispettive presunte preferenze. Si vengono così a produrre perlomeno quattro problemi di non poco conto in riferimento al carattere democratico dell'ordinamento<sup>10</sup>.

È del tutto verosimile, in primo luogo, che i destinatari delle informazioni non abbiano consapevolezza del carattere circoscritto e parziale di quanto viene loro inviato, confidando, anzi, in una presunta oggettività, completezza e neutralità della rete e dei *social*<sup>11</sup>. Anche a motivo della difficoltà di comprendere i complicatissimi passaggi che portano la AI (attraverso le reti neurali artificiali) a elaborare un output sulla base di milioni di input, in secondo luogo, si rende estremamente complicato qualsiasi controllo sulla concreta veridicità e attendibilità delle informazioni fornite. Il fenomeno delle *fake news* pare, in questo senso, difficilmente affrontabile<sup>12</sup>. In termini più ampi, in terzo luogo, lo stesso assetto plurale delle fonti di informazione è sottoposto ad una forte tensione; tensione che porta a indebolire quel *marketplace of ideas* su cui si basa una libera competizione pubblica, il circuito della responsabilità politica e, in ultima analisi, il cuore della logica democratica<sup>13</sup>. Pur evitando catastrofismi, si può segnalare, in quarto luogo, il rischio che un controllo così pervasivo, mirato e individualizzato da parte dei sistemi di AI arrivi a produrre effetti tali da offuscare lo stesso principio di sovranità popolare. Pur continuando ad appartenere formalmente al popolo, la sovranità sarebbe così sostanzialmente esercitata da quanti, per motivi di calcolo politico ed economico, sono in grado di influenzare una quantità di votanti tale da spostare l'esito complessivo della competizione elettorale<sup>14</sup>. Casi come la probabile influenza prodotta da Cambridge Analytica sulle più recenti elezioni presidenziali statunitensi o sulla Brexit (ma la stessa società ha ammesso di aver lavorato in occasione di oltre 200 votazioni sull'intero pianeta) possono ascrivere a tale segnalato rischio<sup>15</sup>; rischio che è in grado, fra l'altro, di produrre

<sup>10</sup> Per i primi tre, cfr. M. FASAN, *Intelligenza artificiale e pluralismo: uso delle tecniche di profilazione nello spazio pubblico democratico*, in *BioLaw Journal – Rivista di BioDiritto*, 1, 2019, 107.

<sup>11</sup> M. AINIS, *Il regno dell'Uroboro. Benvenuti nell'era della solitudine di massa*, Milano, 2018, 11, sottolinea il pericolo costituito dal fatto che, in realtà, «non siamo più liberi di pensare i nostri stessi pensieri [...]. È l'universo autistico in cui siamo rinchiusi anche se per lo più non ci facciamo caso».

<sup>12</sup> In generale, G. PITRUZZELLA, *La libertà di informazione nell'era di Internet*, in *MediaLaws – Rivista di diritto dei media*, 1, 2018, 30 ss. Il fenomeno cd. della *black box*, su cui F. PASQUALE, *The black box society. The secret algorithms that control money and information*, Cambridge-London, 2015, sarà trattato *infra*.

<sup>13</sup> Y. BENCLER ET AL., *Understanding Media and Information Quality in an Age of Artificial Intelligence, Automation, Algorithms and Machine Learning*, in *Berkman Klein Center for Internet and Society, Harvard University*: <https://cyber.harvard.edu/story/2018-07/understanding-media-and-information-quality-age-artificial-intelligence-automation>; D. SHULTZ, *Could Google influence the presidential election?*, in *Science*, Oct. 25, 2016, <https://www.sciencemag.org/news/2016/10/could-google-influence-presidential-election>. Più in generale, Y. BENCLER, R. FARIS, H. ROBERTS, *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*, New York, 2018.

<sup>14</sup> S. ARAL, D. ECKLES, *Protecting elections from social media manipulation*, in *Science*, 365 (6456), 2019, 858.

<sup>15</sup> È del 28 giugno 2019 la notizia della multa di un milione di euro comminata dal Garante per la protezione dei dati personali italiano a Facebook per gli illeciti compiuti nell'ambito del caso Cambridge Analytica. In particolare, si è accertato che «57 italiani avevano scaricato l'app *Thisisyourdigitallife* attraverso la funzione Facebook login e che, in base alla possibilità consentita da questa funzione di condividere i dati degli "amici", l'applicazione aveva poi acquisito i dati di ulteriori 214.077 utenti italiani, senza che questi l'avessero scaricata, fossero stati informati

effetti che (ai sensi dell'art. 48 Cost.) tendono a negare non solo il carattere segreto del voto, ma anche la sua stessa libertà.

Il primo articolo della Costituzione italiana, allora, può costituire la base per tentare di impedire un impiego incontrollato della AI in termini di disinformazione e propaganda politica; un impiego che rischia di generare un modello diffuso di *bubble democracy*<sup>16</sup>, in cui i cittadini, anche in quanto elettori, sono confinati all'interno di sistemi di *social* sempre più chiusi e autoreferenziali, che li portano a ritenere le proprie idee le uniche plausibili, che impediscono un confronto reale e pluralista fra posizioni diverse e che, in ultima analisi, favoriscono una involuzione sociale e politica nel segno della frammentazione, della polarizzazione e della radicalizzazione, con un sostanziale svuotamento dall'interno della sovranità popolare<sup>17</sup>.

#### 4. Il lavoro: potenzialità e rischi (artt. 1 e 4)

L'impatto del sempre più diffuso impiego della AI sul mercato del lavoro costituisce uno degli ambiti di studio più coltivati in questo settore<sup>18</sup>. Pur nella diversità dei campi analizzati e a fronte di cifre e previsioni non coincidenti, la maggior parte delle ricerche indica un iniziale saldo negativo rispetto al numero complessivo di posti di lavoro; saldo che si concentrerà, secondo molti, sulle mansioni più ripetitive, noiose, rischiose o pericolose. Allo stesso modo, in un lasso di tempo che però non sarà breve, si creeranno nuove professionalità legate, ad esempio, al *data mining*, alla programmazione ed al controllo delle modalità di autoapprendimento ed al funzionamento complessivo dell'AI. Molte altre mansioni, da quelle nel mondo medico all'ingegneria, dalla finanza ai trasporti, dalla telecomunicazione all'agricoltura, fino a quelle creative e artistiche subiranno un profondo mutamento e potranno migliorare in termini di rapidità, precisione ed efficienza complessiva<sup>19</sup>.

Un aspetto rilevante, in ogni caso, riguarda la diversa capacità che gli Stati potranno esprimere al fine di approfittare o meno delle opportunità della AI. Anche in questo ambito, infatti, diversi studi hanno dimostrato come la ricchezza e i benefici prodotti dall'innovazione in tale campo non si distribuiranno omogeneamente sul pianeta, ma si concentreranno in quelle economie che avranno adottato strategie di investimento mirate ed efficienti<sup>20</sup>. Per quanto riguarda i lavoratori, così, è prevista, nell'arco dei prossimi dieci anni, una diminuzione di almeno il 10% (dal 40 al 30%) della richiesta di mansioni ripetitive e a basso tasso di abilità digitali; 10% che verrà acquisito dai lavori altamente tecnologizzati (che

---

della cessione dei loro dati e avessero espresso il proprio consenso a questa cessione»: <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9121352>.

<sup>16</sup> D. PALANO, *Verso una bubble-democracy*, in *Formiche*, 2, 2017, 16.

<sup>17</sup> Una discussione tecnica in M. BARHAMGI ET AL., *Social networks data analysis with semantics: application to the radicalization problem*, in *Journal of Ambient Intelligence and Humanized Computing*, 1, 2018.

<sup>18</sup> Fra i più recenti, cfr. *Notes from the frontier: Modeling the impact of AI on the world economy* pubblicato nel settembre 2018 dal *McKinsey Global Institute*, da cui molti degli spunti di seguito riportati: <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-frontier-modeling-the-impact-of-ai-on-the-world-economy>.

<sup>19</sup> In generale, già M. FORD, *Rise of the Robots: Technology and the Threat of Jobless Future*, New York, 2015.

<sup>20</sup> Cfr., fra gli altri, *The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning systems* del 16 maggio 2018: <https://www.accessnow.org/the-toronto-declaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learning-systems/>.

dovrebbero crescere dal 40 al 50%). In termini di reddito, la prima categoria di lavoratori potrà affrontare una stagnazione o una diminuzione complessiva, a vantaggio di un incremento della seconda. Se tali mutamenti non sono certo nuovi nel mondo del lavoro, avendo contraddistinto ogni precedente ondata di forte innovazione, la AI si caratterizza per la velocità con cui tali mutamenti avverranno; velocità che renderà molto complessa e praticamente impossibile, ad esempio, la conversione degli stessi lavoratori da una mansione non specializzata ad una fortemente specializzata. Su queste basi, si è calcolato che i Paesi che avranno la volontà politica e le capacità tecniche di investire in termini strutturali nella AI potranno vedere una crescita economica aggiuntiva stimata fra il 20 e il 25%; cifra che sarà irrecuperabilmente persa, forse in termini ancora più incisivi, dagli Stati che non avranno adottato simili strategie.

Quanto pare prospettarsi, insomma, è l'emersione di un forte rischio di *AI divide*, in termini di aumento della (già ampia) distanza fra Paesi, società e categorie di lavoratori – per così dire – *AI friendly* e non<sup>21</sup>. Anche in questa prospettiva, la Costituzione italiana contiene indicazioni che dovrebbero orientare la politica a tenere conto della AI in termini globali, al fine, ad esempio, di arricchire i percorsi scolastici e universitari e di adeguare la formazione professionale (art. 35), di rendere il mercato del lavoro più equilibrato e pronto per le sfide tecnologiche, di preparare la società nel suo complesso a gestire consapevolmente un settore carico di potenzialità ma anche di rischi. Parafrasando l'art. 4 Cost., insomma, pare importante rivolgere un appello forte a che si promuovano, con strumenti aggiornati ed efficaci, le condizioni che potranno rendere effettivo il diritto al lavoro in una società satura di AI. Solo con questo impegno, da avviare in tempi rapidissimi, la Repubblica potrà, secondo l'imperativo dell'art. 1 e nel solco del principio lavorista, ritrovarsi fondata su un lavoro che non crei nuove e più profonde discriminazioni, ma si riveli vero motore di mobilità e arricchimento sociale e di realizzazione personale<sup>22</sup>.

### 5. AI e eguaglianza (art. 3): salute (art. 32) e giustizia (art. 111)

Le tecniche di AI possono essere utilmente impiegate per superare diseguaglianze di fatto e di diritto, ma possono anche nascondere gravi forme di discriminazione. Vista l'ampiezza del tema, ne tratterò, in questa sede, intrecciando due soli ambiti in qualche modo paradigmatici dei vantaggi e dei possibili rischi: quello riferito al diritto alla salute e quello relativo all'amministrazione della giustizia.

Nell'ambito della salute, la AI può essere efficacemente impiegata allo scopo di migliorare le tecniche di intervento piuttosto che a quello di permettere ad ognuno di sviluppare le proprie capacità, superando disabilità e svantaggi di diverso genere. Su questa linea, così, sono moltissimi gli ambiti che possono guadagnare in termini di maggior efficacia, velocità e accuratezza nel curare la salute delle persone: dall'assistenza agli anziani o ai bambini autistici alla diagnostica basata su *brain imaging*, dalla chirurgia robotica alla sperimentazione, dalla medicina di precisione alle protesi<sup>23</sup>. In urgenza, per fare

<sup>21</sup> Fra gli altri, A. LÓPEZ PELÁEZ (ed.), *The Robotics Divide. A New Frontier in the 21st Century?*, London, 2014.

<sup>22</sup> Su cui già C. MORTATI, *Costituzione dello Stato*, II, *La costituzione italiana*, in *Enciclopedia di diritto*, vol. XI, Milano, 1962, 14.

<sup>23</sup> UNESCO, *World Commission on the Ethics of Scientific Knowledge and Technology (COMEST), Report on Robotics Ethics*, Paris, 14 September 2017, 30. In generale, si vedano gli articoli pubblicati nella rivista specializzata *Artificial Intelligence in Medicine*: <https://www.journals.elsevier.com/artificial-intelligence-in-medicine>.

un esempio su cui vi è ormai molta letteratura, l'impiego di algoritmi permette di giungere ad ipotesi di diagnosi in tempi molto rapidi, confrontando simultaneamente migliaia di casi; in questo modo, la AI permette interventi rapidi e più efficaci sul malato, rendendo anche più celeri i flussi di pazienti in pronto soccorso<sup>24</sup>. E nel settore chirurgico, l'impiego di macchinari robotizzati guidati anche dalla AI permette interventi complicatissimi, assicurando minore invasività, maggior precisione e accuratezza<sup>25</sup>.

Nell'ambito della giustizia, in prospettiva equivalente, si sono proposte e talvolta impiegate tecniche di decisione assistite o addirittura affidate alla AI<sup>26</sup>. A titolo di esempio, uno studio statunitense retrospettivo ha permesso di dimostrare come, con l'impiego della AI, si sarebbero potute adottare decisioni di rilascio su cauzione più mirate. In particolare, si sarebbero potuti scarcerare molti più indagati (il 25%), senza per questo far crescere i tassi di criminalità, oppure, a parità di numero di rilasci, si sarebbe potuto fra diminuire di oltre il 40% l'indice di criminalità<sup>27</sup>. E l'algoritmo avrebbe anche ridotto le disparità di trattamento legate all'origine razziale delle persone coinvolte.

I dati complessivi al riguardo, tuttavia, sono discordanti e alcune ricerche hanno dimostrato come la AI possa funzionare sulla base degli stessi pregiudizi umani. Così, per motivi che vanno dall'incidenza statistica di alcuni fattori oggettivi fino ai criteri soggettivi individuati dai programmatori, si è rilevato come il colore della pelle abbia giocato in alcuni casi di applicazione algoritmica un ruolo rilevante nel valutare il rischio di pericolosità sociale degli imputati<sup>28</sup>.

Nell'elaborare le proprie decisioni, inoltre, i sistemi informatici di AI, applicati alla giustizia come alla medicina, soffrono di un difetto di trasparenza che rischia di pregiudicarne all'origine l'utilizzo. Si tratta del cd. fenomeno della *black box*, secondo cui, con l'impiego soprattutto del *machine e deep learning*, non è dato comprendere, nemmeno ai programmatori, il percorso attraverso cui la macchina produce il risultato della propria attività<sup>29</sup>. Tale profilo è particolarmente grave in quanto non permette di cogliere, né quindi di controllare, i passaggi alla base dell'esito raggiunto dal sistema. In quest'ottica, qualsiasi decisione giudiziaria affidata alla AI sarebbe attaccabile in quanto non provvista della

<sup>24</sup> N. LIU ET AL., *Artificial intelligence in emergency medicine*, in *Journal of Emergency and Critical Care Medicine*, 2, 2018, 82, individuano i tre settori di principale utilizzo dell'AI nella medicina d'emergenza: «AI in predictive modeling, AI in patient monitoring, and AI in emergency department operations».

<sup>25</sup> D. HASHIMOTO ET AL., *Artificial Intelligence in Surgery: Promises and Perils*, in *Annals of Surgery*, 268(1), 2018, 70.

<sup>26</sup> In generale, cfr. A. GARAPON, J. LASSÈGUE, *Justice digitale*, Paris, 2018.

<sup>27</sup> Si tratta di uno studio condotto sull'analisi retrospettiva di oltre 758.000 casi di decisioni relative al rilascio su cauzione adottate fra il 2008 e i 2013 nel distretto di New York. L'utilizzo di algoritmi di AI avrebbe potuto permettere la riduzione dell'indice di criminalità di quasi il 25% mantenendo inalterato il tasso di detenzione. Altrimenti, avrebbe potuto condurre ad una riduzione di oltre il 40% del tasso di carcerazione senza per questo aumentare i livelli di criminalità. J. KLEINBERG ET AL., *Human Decisions and Machine Predictions*, in *The Quarterly Journal of Economics*, 133(1), 1 February 2018, 241.

<sup>28</sup> Si veda il caso *Compas* ed altri riportati da A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, *BioLaw Journal – Rivista di BioDiritto*, 1, 2019; J. ZOU, L. SCHIEBINGER, *AI can be sexist and racist – it's time to make it fair*, in *Nature*, 559, 12 July 2018, 324; M. SUN, M. GERCHICK, *The Scales of (Algorithmic) Justice: Tradeoffs and Remedies*, in *AI Matters*, 2, 2019, 30; L. GIACOMELLI, *Big brother is «gendering» you*, *BioLaw Journal – Rivista di BioDiritto*, 2, 2019, 284.

<sup>29</sup> Si veda il capitol dedicato a *Inside the Black Box: Understanding AI Decision Making*, in S. MOHANTY, S. VYAS, *How to Compete in the Age of Artificial Intelligence*, Berkeley, 2018, 91.

necessaria motivazione (art. 111 Cost.), rendendo sostanzialmente incostituzionale ogni sistema di giustizia automatizzata. E se l'atto medico non gode della stessa garanzia costituzionale, è possibile dubitare della legittimazione complessiva di una medicina che basi le proprie decisioni su operazioni di carattere informatico di cui nessuno può garantire comprensione e controllo.

Altri profili possono condurre ad un accostamento fra medicina e giustizia teso a far emergere vantaggi e rischi dell'impiego dell'AI. Una sentenza frutto di un algoritmo programmato ed applicato in maniera corretta, ad esempio, potrebbe incarnare i requisiti di terzietà, imparzialità e predeterminazione, facendo paradossalmente assumere al giudice artificiale le vesti del perfetto giudice naturale (art. 25). E la ragionevole durata del processo sarebbe garantita fin da principio. Allo stesso tempo, però, ci si potrebbe interrogare sulla possibilità di ricondurre alla funzione giurisdizionale un'attività che funzionerebbe in termini prevalentemente probabilistici. La *Predictive Justice*, in questo senso, sarebbe in quanto tale *Justice*? In prospettiva più puntuale, la giustizia algoritmica potrebbe revocare in dubbio l'utilità dell'appello o del ricorso in Cassazione (verso un computer più potente? Verso una sezione composta, questa volta, da soli umani?) come del rivendicato (già nella *Magna Carta*) diritto ad un giudizio da parte dei propri pari; pari cui, in quanto tale, non potrebbe condursi nemmeno la più umanizzata delle macchine dotate di AI<sup>30</sup>. Alla luce di rischi di questo genere, può indicarsi la legge di revisione della giustizia francese del marzo 2019 il cui art. 33 vieta l'utilizzo di strumenti destinati alla giustizia predittiva: «L'identità dei magistrati e dei funzionari del registro non possono essere utilizzati allo scopo di valutare, analizzare, confrontare o prevedere le loro pratiche professionali effettive o presunte»<sup>31</sup>.

In medicina, in termini simili, un dottore-robot potrebbe analizzare in tempo reale milioni di dati e informazioni cliniche, prospettando diagnosi e terapie basate sul confronto con moltissimi casi affini. Potrebbe trattare tutti i pazienti allo stesso modo senza sentire stanchezza o avere stringenti limiti di tempo. Un eguale distacco verso i malati, tuttavia, sarebbe difficilmente compatibile con le più recenti acquisizioni, anche legislative, in tema di umanizzazione della medicina, che hanno portato a qualificare la relazione medico-paziente in termini di cura e di fiducia<sup>32</sup>.

A fronte di tale quantità di possibili vantaggi e rischi, si conferma, anche in questi settori, l'esigenza che siano i principi costituzionali ricordati ad orientare la disciplina della AI e del suo utilizzo.

## 6. AI e privacy: il rischio per antonomasia

Proseguendo nell'analisi dei principi costituzionali che possono fungere da utile orientamento per la AI, un passaggio d'obbligo è costituito dalla cd. privacy. Riprendendo quanto detto in apertura in riferimento alla mole di dati in circolazione (44 zettabyte), può richiamarsi come le nostre attività

<sup>30</sup> Se ogni riferimento alla *reflective judiciary* (S. TURENNE (ed.), *Fair Reflection of Society in Judicial Systems - A Comparative Study*, Cham, 2015) sarebbe improponibile, può ricordarsi come alcuni progetti, come AI4ALL, abbiano richiamato l'attenzione sulla necessità di ampliare la platea dei programmatori a donne, persone di colore e, in generale, alle categorie sociali meno rappresentate, come strumento per evitare che il funzionamento dei sistemi informatici sia viziato dai vari pregiudizi di cui soffrono tipicamente i maschi Wasp: <http://ai-4-all.org/>.

<sup>31</sup> LOI n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice. Il divieto è assistito da una pesante sanzione penale.

<sup>32</sup> Il riferimento va al primo articolo della legge n. 219 del 2017.

quotidiane e l'ambiente che ci circonda presentino una infinità di occasioni per carpire e diffondere dati personali. L'azione combinata di ogni nostra ricerca sul computer o click sul telefonino con i pagamenti con carte di credito o debito, con i rapporti con la pubblica amministrazione, ma anche con ogni accensione della televisione e dei diversi dispositivi riconducibili alla IoT o con la circolazione nelle cd. *smart cities* crea una quantità ed una qualità di informazioni riconducibili ad ogni singola persona tali da poterne tracciare un profilo dettagliato, comprensivo, ad esempio, dello stato di salute, del livello culturale ed economico, dei gusti alimentari, dell'orientamento sessuale e delle varie preferenze, fino all'affidabilità finanziaria o alla propensione a delinquere<sup>33</sup>. Grazie all'utilizzo dell'AI, infatti, un volume di dati altrimenti ingestibile, può essere trattato con rapidità e precisione al fine di effettuare una profilazione individuale molto accurata<sup>34</sup>. Si rende così evidente il rischio generalizzato di una diffusa e capillare intrusione nella sfera più intima di ognuno, con il rischio di una palese violazione del diritto alla riservatezza e di una esposizione di dati anche personali che potranno essere in futuro utilizzati in qualsiasi occasione, dalla richiesta di un mutuo ad un colloquio di lavoro fino alla determinazione del rischio di pericolosità sociale<sup>35</sup>. E di fronte a tali rischi, può segnalarsi la decisione del *Board of Supervisors* della città di San Francisco del maggio 2019 che ha deciso di vietare l'utilizzo di software di riconoscimento facciale anche da parte della polizia<sup>36</sup>.

In quanto dato costituzionale, il diritto alla riservatezza ha visto in origine una tutela ottenuta attraverso lo strumento del consenso informato. Nel corso del tempo, tuttavia, tale categoria si è dimostrata inefficace nel proteggere la persona nei confronti delle applicazioni delle tecnologie dell'informazione e della comunicazione (ICT), convertendosi da strumento a tutela della riservatezza della persona a mezzo attraverso il quale quotidianamente esponiamo i nostri dati a chiunque ci fornisca un servizio attraverso internet<sup>37</sup>. È dato evidente, ad esempio, come ognuno di noi accetti le condizioni contrattuali legate alle varie applicazioni per i cellulari o per i computer senza nemmeno considerare la lettura delle stesse. In questi termini, il collegamento fra la necessaria informazione e la scelta da attuare viene solitamente reciso, portando alla diffusione di una forma di consenso consapevolmente disinformato. Si tratta – se vogliamo – di una peculiare forma di eterogenesi dei fini, in cui un istituto nato per tutelare la dimensione personale è divenuto una finzione e, anzi, si è trasformato nel mezzo più efficace per penetrarla e per esporci quotidianamente alla profilazione di ogni nostra dimensione e attività.

<sup>33</sup> In generale, F. PIZZETTI (a cura di), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018.

<sup>34</sup> In generale, B.C. NEWELL, T. TIMAN, B.J. KOOPS (eds.), *Surveillance, Privacy, and Public Space*, London, 2017.

<sup>35</sup> La letteratura è sul punto sterminata. In termini generali, S. GUTWIRTH, P. DE HERT, R. LEENES, *Data protection on the Move*, Dordrecht, 2016. Su un utilizzo particolarmente inquietante dei dati da parte del governo cinese, K. ROTH, M. WANG, *Data Leviathan: China's Burgeoning Surveillance State*, in *The New York Review of Books*, August 16, 2019.

<sup>36</sup> Ne ha dato notizia il *New York Times* del 14 maggio 2019. Cfr. Anche *San Francisco residents question city's facial recognition ban*, in *Biometric Technology Today*, 6, 2019, 1.

<sup>37</sup> Cfr. in generale, J.E. COHEN, *What privacy is for*, in *Harvard Law Review*, 126, 2012, 1904. Il regolamento (UE) 2016/679 (il cd. GDPR: *General Data Protection Regulation*) considera ancora il consenso informato quale pietra angolare della tutela dei dati personali, dimostrandosi così un'arma del tutto spenta rispetto alle innovazioni nel campo dell'ICT ottenute attraverso l'AI.

A fronte di questa situazione, si potrebbe, senza qualche difficoltà, abbandonare l'impostazione individualista del consenso e rinforzarlo attraverso una garanzia posta (anche) a livello statale. Senza arrivare a posizioni apertamente paternaliste, si potrebbe fare riferimento al secondo comma dell'art. 41 Cost., imponendo alle imprese che gestiscono i dati limiti corrispondenti all'utilità sociale e al rispetto della sicurezza, della libertà e della dignità umana.

## 7. La decisione automatizzata

Questa breve rassegna di principi costituzionali in grado di orientare la AI ed il suo utilizzo potrebbe proseguire. Si pensi, ad esempio, alla cd. agricoltura di precisione, in cui, attraverso l'analisi dei dati e l'impiego di droni si potrebbe passare da un impiego generalizzato e preventivo di pesticidi all'individuazione delle singole piante da trattare e ad interventi specifici e mirati. Tale strategia permetterebbe di ridurre in termini considerevoli l'utilizzo di sostanze chimiche a beneficio della tutela ambientale (art. 9, secondo comma, Cost.) e della salute collettiva (art. 32). Un altro esempio può riferirsi all'impiego dei droni in scenari bellici che, attraverso tecniche di riconoscimento facciale e di bombardamento mirato, sono utilizzati per eliminare terroristi o presunti tali. Al di là delle possibili imprecisioni nelle tecniche utilizzate, tale strategia potrebbe trovare, in Italia, un ostacolo generalizzato nel proclamato ripudio della guerra come «strumento di offesa alla libertà degli altri popoli e come mezzo di risoluzione delle controversie internazionali» (art. 11 Cost.).

A fronte della molteplicità dei possibili rilievi costituzionali in tema di AI, quanto qui interessa porre in evidenza è una ultima questione, a mio avviso centrale: quella riferibile alla possibilità di assumere decisioni che incidono significativamente su una persona e sui suoi diritti grazie all'assistenza di sistemi di AI o addirittura attraverso la sostituzione della componente umana con quella tecnologica<sup>38</sup>.

### 7.1. Sostituzione v. assistenza

Un giudizio d'insieme e generalizzato sull'opportunità di una sostituzione della decisione umana con quella automatizzata non è agevole. Come visto, ci sono ambiti in cui paiono prevalere i vantaggi ed altri in cui sono i rischi ad imporsi. Alcuni autori, tuttavia, sostengono una superiorità complessiva dell'AI rispetto all'intelletto umano. A causa del rumore di fondo che caratterizzerebbe quest'ultimo, dovuto a componenti riconducibili alla nostra storia personale, agli affetti o alle persone di cui ci fidiamo, sarebbe così consigliabile, in un futuro prossimo in cui la AI potrà gestire una sufficiente mole di dati, far prendere le decisioni più rilevanti ad essa. Lo psicologo Daniel Kahneman, considerato uno dei padri della economia comportamentale e premio Nobel nel 2002, sostiene l'impossibilità di esprimere scelte umane realmente razionali<sup>39</sup>. In una sorta di precomprensione totale, gli uomini assumerebbero decisioni del tutto illogiche e insensate, e solo in un secondo tempo creerebbero argomenti

<sup>38</sup> In questa parte del lavoro, riprendo, sintetizzandole ed aggiornandole, considerazioni già svolte in *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, cit.

<sup>39</sup> Si vedano le posizioni dell'autore espresse al *World Business Forum* di New York nel novembre 2018. In Italia, sono state riprese da *La Stampa* del 16 novembre.

falsamente ragionevoli per supportarle. Su questa linea, Kahneman ha sostenuto che la AI «will make better choices for the individual because it will be noise free»<sup>40</sup>.

A fronte di questa impostazione, ritengo che l'intelligenza umana possa essere utilmente supportata, ma non debba essere sostituita integralmente da quella artificiale. I motivi alla base di questa opinione sono di natura sia tecnica, sia filosofico-antropologica più ampia.

a) Da un punto di vista tecnico, in primo luogo, si registra come le macchine dotate di AI siano insuperabili nello svolgere alcune funzioni (*weak AI*), ma siano ben lontane dal replicare l'insieme complessivo delle abilità e dei comportamenti umani (*strong AI*). Inoltre, anche nello svolgere specifici compiti come il riconoscimento di visi o di immagini i sistemi di AI hanno dimostrato di incorrere in errori che sarebbero evidenti ad ogni persona<sup>41</sup>.

b) In secondo luogo, sempre da un punto di vista tecnico, non pare potersi sostenere che le logiche algoritmiche siano di per sé più neutrali ed oggettive dei ragionamenti umani che pure stanno alla base della loro costruzione, programmazione e funzionamento. Tanto la scelta degli algoritmi con cui equipaggiare la macchina, quanto la selezione degli scenari e degli esempi con cui "nutrirla", ad esempio, sono operazioni affidate all'operatore umano<sup>42</sup>. In questa logica, si può affermare che il pensiero artificiale è comunque, perlomeno nelle sue condizioni iniziali di base, influenzato a monte dal pensiero umano e dalla sua variabilità, incostanza e discrezionalità.

c) In terzo luogo, il segnalato fenomeno della *black box* si pone come ostacolo forte, di natura tecnica, alla possibilità di tracciare i passaggi logico-argomentativi che permettono alla macchina di decidere in autonomia: se l'output finale del procedimento è noto, non lo è l'iter che lo ha generato. Stando così le cose, non pare possibile verificare, anche da un punto di vista sperimentale, la congruità dei passaggi svolti e quindi la correttezza dell'attività interna alla base della decisione.

A fronte di tali tre obiezioni tecniche all'ipotesi della sostituzione, si potrebbe dare una risposta basata sul fatto che nel tempo si troveranno gli strumenti per (a) superare la limitazione nelle capacità dell'AI, (b) la sua non neutralità e (c) opacità. Non può infatti escludersi che, in futuro, la *strong AI* diventi realtà (a); che la AI diventi responsabile della propria stessa programmazione (b); che i progressi tecnologici possano rendere comprensibili le operazioni interne al *machine learning* (c).

<sup>40</sup> Cfr. la conferenza *The Economics of AI* (13-14 settembre 2017, Toronto). La dichiarazione è attribuita a Kahneman da J. PETHOKOUKIS, *Nobel laureate Daniel Kahneman on AI: 'It's very difficult to imagine that with sufficient data there will remain things that only humans can do'*, nel sito <http://www.aei.org/publication/nobel-laureate-daniel-kahneman-on-a-i-its-very-difficult-to-imagine-that-with-sufficient-data-there-will-remain-things-that-only-humans-can-do/>. Cfr. Anche J.N. MATHIAS, *Bias and Noise: Daniel Kahneman on Errors in Decision-Making*, in *Medium*, 17 ottobre 2017: <https://medium.com/@natematias/bias-and-noise-daniel-kahneman-on-errors-in-decision-making-6bc844ff5194>.

<sup>41</sup> Si è dimostrato, ad esempio, come l'accuratezza del riconoscimento facciale dipenda da variabili strettamente legate alle caratteristiche etnico-razziali e sessuali: I. RAJI, J. BUOLAMWINI, *Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products*. Conference on Artificial Intelligence, Ethics, and Society, pubblicato nel sito del MIT Media Lab il 24 gennaio 2019: <https://www.media.mit.edu/publications/actionable-auditing-investigating-the-impact-of-publicly-naming-biased-performance-results-of-commercial-ai-products/>. In generale, A. ROSENFELD, R. ZEMEL, J.K. TSOTSOS, *The Elephant in the Room*, 9 agosto 2018, reperibile al sito della Cornell University: <https://arxiv.org/abs/1808.03305>.

<sup>42</sup> In generale, G-Z. YANG, P. DARIO, D. KRAGIC, *Social robotics—Trust, learning, and social interaction*, in *Science Robotics*, 3(21), 2018.

In ogni caso, tuttavia, mi pare rimangano alcune questioni non superabili di natura filosofico-antropologica. Da più parti ad esempio, si sostiene che alcune qualità umane non potranno mai essere sostituite da componenti artificiali: si tratta, a seconda delle impostazioni, dell'immaginazione; della capacità di dare vita a processi creativi; della coscienza, intesa secondo la teoria dell'informazione integrata; di creatività e ispirazione, empatia ed emozioni<sup>43</sup>. E anche il beneficio del dubbio, con il correlato senso di curiosità, e il "sapere di non sapere" sono caratteristiche che contraddistinguono l'umano e la sua ricerca di senso.

L'ipotesi più convincente, quindi, riguarda non la completa sostituzione dell'umano con la AI, ma un'attività di assistenza: una collaborazione in termini di sussidiarietà, in cui ogni parte svolge i compiti che maggiormente le si addicono. Per quanto riguarda l'amministrazione della Giustizia, così, si sono sottolineati i possibili vantaggi del supporto algoritmico in termini, ad esempio, di rapidità e ampiezza dell'istruttoria; di completezza degli elementi su cui costruire la decisione; di certezza (o esattezza) del diritto<sup>44</sup>. E nell'ambito medico – per mantenere la simmetria utilizzata in precedenza – l'esame umano di traccati e risonanze potrebbe certamente avvalersi di elaborazioni statistiche che solo la AI può fare sulla base di quantità di dati e di raffronti altrimenti ingestibili, come il ricercatore che lavora sulla medicina di precisione non potrebbe concretamente svolgere i propri studi senza l'ausilio della AI.

Su queste basi, si possono fissare almeno tre profili all'interno dei quali proporre nuove (o rinnovate) posizioni giuridiche che permettano di favorire una relazione uomo-AI equilibrata e costituzionalmente orientata.

## 8. I diritti per una AI costituzionalmente orientata

a) Un primo diritto che si dovrebbe poter invocare di fronte all'utilizzo della AI è quello di ottenere una spiegazione dei passaggi attraverso i quali la macchina ha generato il proprio risultato. Non si tratterebbe di una illustrazione tecnico-informatica, inaccessibile ai più, ma di un chiarimento dato in termini comprensibili delle logiche che hanno permesso alla macchina di esercitare l'autonomia alla base della propria decisione. In questo modo, si estenderebbe il diritto alla motivazione, espressamente previsto in Costituzione (111, sesto comma) per i provvedimenti giurisdizionali e dalla legge 241 del 1990 (art. 3) per quelli amministrativi, ad ogni decisione, pubblica o privata, che possa incidere sulla posizione giuridica dell'interessato<sup>45</sup>.

<sup>43</sup> Si tratta, rispettivamente, di: Peter Ware Higgs, premio Nobel per la Fisica nel 2013; Marc Mézard, fisico, direttore della *École Normale Supérieure* di Parigi; Giulio Tononi, psichiatra e neuroscienziato italiano, direttore del Center for Sleep and Consciousness dell'Università del Wisconsin; Roberto Cingolani, fisico, direttore scientifico dell'Istituto Italiano di Tecnologia (IIT) di Genova.

<sup>44</sup> M. LUCIANI, *La decisione giudiziaria robotica*, in *Rivista AIC*, 3, 2018, 878 e C. PINELLI, *Certezza del diritto e compiti dei giuristi*, in *Diritto pubblico*, 2, 2019, 556; C. BARBARO, Y. MENECEUR, *Issues in the use of artificial intelligence (AI) algorithms in judicial systems*, in *European Commission for the Efficiency of Justice Newsletter*, Council of Europe, n. 16, August 2018, 3.

<sup>45</sup> Sull'utilizzo della AI da parte della pubblica amministrazione, fra gli altri, S. CIVITARESE MATTEUCCI, *Umano troppo umano». Decisioni amministrative automatizzate e principio di legalità*; F. COSTANTINO, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, entrambi in *Diritto pubblico*, 1, 2019, 5 e 43; I.A. NICOTRA, V. VARONE, *L'algoritmo, intelligente ma non troppo*, in *Rivista AIC*, 4, 2019, 86.

Tale diritto, ad oggi, imporrebbe alla tecnologia informatica un onere che alcuni considerano tecnicamente insostenibile o impossibile da reggere. A fronte della sostanziale mancanza di legittimazione di ogni decisione di cui non si possano comprendere le motivazioni, pure, si sono avviati numerosi studi orientati, da un lato, a rendere trasparenti le logiche del *machine learning* e, dall'altro, a superare tali tecniche e ad adottare sistemi differenti che possano garantire una maggiore trasparenza<sup>46</sup>. In ogni caso, la proposta di un diritto così significativo e pregnante come quello alla motivazione non pare poter essere ostacolato a motivo delle difficoltà tecniche legate al suo rispetto.

b) Una seconda posizione giuridica soggettiva cui si potrebbe attribuire rilievo sostanzialmente costituzionale consiste nel diritto ad essere resi consapevoli della natura, umana o artificiale, del proprio interlocutore<sup>47</sup>. A fronte della possibilità tecnica di creare sistemi artificiali che possano essere confusi con persone umane (su questo si basava già il test di Turing)<sup>48</sup>, pare essenziale confermare tale diritto per una serie di motivi. Fra questi, si può richiamare il diritto ad essere correttamente informati riguardo alle possibili conseguenze delle attività che conduciamo nella nostra vita quotidiana. Una macchina potrebbe immettere all'istante in un circuito potenzialmente globale e incontrollato (il *cloud*) tutti i dati che ad essa trasmettiamo; non sapendo se stiamo colloquiando con un nostro simile o con una macchina, rischieremo di perdere definitivamente il controllo su di essi. L'impiego di robot antropomorfi in taluni settori, inoltre, solleva questioni di particolare rilievo. Potremmo sentirci raggirati o traditi, ad esempio, nel momento in cui consegniamo la nostra salute, le nostre confidenze o i nostri affetti a robot antropomorfi che riteniamo essere umani. E l'ignoranza sulla natura del nostro interlocutore potrebbe generare fraintendimenti, falsi affidamenti o tradire eventuali aspettative di comprensione empatica.

In ultima analisi, la non conoscenza del carattere umano o artificiale della nostra controparte, perlomeno nello svolgimento di alcune attività, potrebbe avere ripercussioni sulla nostra dignità. Come affermato in un rapporto dell'Unesco, «Dignity is inherent to human beings, not to machines or robots. Therefore, robots and humans are not to be confused even if an android robot has the seductive appearance of a human, or if a powerful cognitive robot has learning capacity that exceeds individual human cognition»<sup>49</sup>.

c) La terza posizione da proporre nella relazione uomo-AI consiste nel diritto ad essere destinatari di decisioni che siano il risultato di un processo in cui sia presente una significativa componente umana<sup>50</sup>. Tale posizione si collega, in primo luogo, alla opportunità di ricevere provvedimenti che in qualche modo siano espressione di nostri pari; opportunità che per alcune decisioni non trascurabili, come per

<sup>46</sup> C. RUDIN, *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead*, in *Nature Machine Intelligence*, 1, 2019, 206.

<sup>47</sup> Sull'opportunità che «people have a right to know whether they are dealing with a human being or with an AI artefact» si interroga lo *Statement on Artificial Intelligence, Robotics and "Autonomous Systems"*, dello *European Group on Ethics in Science and New Technologies (EGE)*, Brussels, 9 March 2018, 11.

<sup>48</sup> Si tratta del famoso test in grado di rilevare, negli intendimenti dell'autore, se una macchina è in grado di pensare, di farlo autonomamente e di farlo come un essere umano: A.M. TURING, *Computing Machinery and Intelligence*, in *Mind*, 59, 1950, 433 ss.

<sup>49</sup> *Report of COMEST on Robotics Ethics*, cit., 50.

<sup>50</sup> Cfr. A. PAJNO ET AL., *AI: profili giuridici. Intelligenza Artificiale: criticità emergenti e sfide per il giurista*, in *BioLaw Journal – Rivista di BioDiritto*, 3, 2019, 205 ss.; A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., 56, ne parla in termini di principio di non esclusività.

quelle giurisdizionali, è già riconosciuta attraverso una interpretazione dei dati giuridici esistenti (il tradizionale diritto ad una *jury of peers*). Il diritto ad una decisione almeno in parte umana può collegarsi, in secondo luogo, alle segnalate possibilità di errore e pregiudizio che, almeno per ora, contraddistinguono le operazioni algoritmiche alla base della AI. Nella stessa ottica, in terzo luogo, il fenomeno della *black box* e la mancanza di trasparenza ad esso collegata milita a favore di una supervisione umana del procedimento<sup>51</sup>. E questo, non solo per i provvedimenti di carattere giuridico (amministrativi o giurisdizionali) ma per ogni decisione pubblica o privata la cui rilevanza e le cui ricadute concrete sulle persone impongano l'opportunità di un vaglio dell'apparato motivazionale. In questi termini, da ultimo, il diritto ad una decisione non interamente automatizzata si basa anche sulla necessità di individuare la titolarità della funzione esercitata e la rispettiva responsabilità complessiva<sup>52</sup>.

Alcuni segnali nella direzione di un riconoscimento della posizione in parola vengono dal diritto dell'UE e, in particolare dal citato GDPR (regolamento 2016/679 *General Data Protection Regulation*). Al considerando 71, tale atto indica che l'interessato «dovrebbe avere il diritto di non essere sottoposto a una decisione, che possa includere una misura che valuti aspetti personali che lo riguardano, che sia basata unicamente su un trattamento automatizzato e che produca effetti giuridici che lo riguardano o incida in modo analogo significativamente sulla sua persona, quali il rifiuto automatico di una domanda di credito online o pratiche di assunzione elettronica senza interventi umani».<sup>53</sup> E l'art. 22 dello stesso regolamento (Processo decisionale automatizzato relativo alle persone fisiche) inizia con «il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona»<sup>54</sup>. La bontà di tali disposizioni rischia peraltro di essere pregiudicata a motivo delle eccezioni che lo stesso regolamento prevede nei casi in cui la decisione «a) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento; [...] c) si basi sul consenso esplicito dell'interessato». Tale ultima clausola, in particolare, permette di svuotare di tutela sostanziale gran parte del diritto a una decisione non completamente automatizzata, visto la logica "disinformata" con cui – come accennato – gestiamo regolarmente i nostri consensi.

<sup>51</sup> A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., ne parla in termini di principi di conoscibilità e di comprensibilità.

<sup>52</sup> Nell'ambito del pubblico, perlomeno, i funzionari e i dipendenti dello Stato sono direttamente responsabili degli atti compiuti in violazione dei diritti (art. 28 Cost.).

<sup>53</sup> Il considerando prosegue indicando che tale trattamento «comprende la 'profilazione', che consiste in una forma di trattamento automatizzato dei dati personali che valuta aspetti personali concernenti una persona fisica, in particolare al fine di analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze o gli interessi personali, l'affidabilità o il comportamento, l'ubicazione o gli spostamenti dell'interessato, ove ciò produca effetti giuridici che la riguardano o incida in modo analogo significativamente sulla sua persona».

<sup>54</sup> Un commento in L.A. BYGRAVE, *EU data protection law falls short as desirable model for algorithmic regulation*, in L. ANDREWS ET AL., *Algorithmic Regulation*, cit., 31. Tale disposizione trova un diretto precedente nell'art. 15 della (ora abrogata) direttiva europea n. 95/46/CE relativa alla tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati.

## 9. Conclusioni: le potenzialità della Costituzione italiana nell'orientare una disciplina della AI

La Costituzione italiana affida alla Repubblica il compito di promuovere la ricerca scientifica e tecnica (art. 9), ricordando come la scienza debba procedere in un regime di libertà (art. 33). Le sue applicazioni, quando siano fatte proprie dall'iniziativa economica privata, non possono tuttavia contrastare l'utilità sociale né recare danno alla sicurezza, alla libertà e alla dignità umana. Che sia in mano privata o pubblica, inoltre, l'attività economica può essere indirizzata e coordinata a fini sociali (art. 41).

Tali principi possono costituire un efficace cornice entro cui inserire una regolamentazione della AI costituzionalmente orientata, che la indirizzi verso scopi di progresso scientifico, economico e sociale, oltre che di generale benessere.

Per raggiungere tale risultato, tuttavia, è necessario che la riflessione etica e giuridica e che la costruzione di una disciplina adeguata procedano in tempi rapidi, di pari passo con la velocità dei progressi scientifici. È essenziale, infatti, che il diritto non inseguia le applicazioni della AI, ma che intervenga a monte, ponendo principi e regole *by design* – per così dire.

Tale esigenza è presente a molti ricercatori che hanno indicato come la AI possa trovare una legittimazione complessiva solo a condizione di un suo utilizzo adeguatamente disciplinato; e sulla stessa linea, si sono mosse alcune delle società che maggiormente fanno ricerca e impiegano tale tecnologia<sup>55</sup>.

Le sfide da affrontare, però, sono molte e particolarmente impegnative. Esse riguardano, per fare qualche esempio, la capacità da parte statale di regolare settori economici in cui le parti private svolgono un ruolo dominante; la possibilità di agire a livello non solo nazionale o internazionale, ma globale, vista la mobilità del *cloud*; l'abilità nel catalizzare una condivisione d'insieme su principi, diritti e doveri di nuova generazione; l'attitudine a rendere consapevole e avvertita la popolazione, a partire dalle generazioni più giovani, delle potenzialità e dei rischi della AI<sup>56</sup>.

Si delinea, insomma, una nuova fase di riflessione sui poteri pubblici e privati e sul ruolo del diritto e dei diritti. Si tratta di una fase che, nel guardare alla disciplina dell'artificiale, dovrà anzitutto guardare al ruolo e alle prerogative che vorremo riservare all'umano; una fase in cui – come illustrato – la Costituzione italiana può dare una molteplicità di spunti di forte interesse.

<sup>55</sup> Cfr., ad esempio, J. HEMPEL, *Fei-Fei Li's Quest to Make AI Better for Humanity*, in *Wired*, 13 novembre 2018, (<https://www.wired.com/story/fei-fei-li-artificial-intelligence-humanity/>).

<sup>56</sup> Spunti in prospettiva bioetica in S. QUINTARELLI ET AL., *AI: profili etici. Una prospettiva etica sull'Intelligenza Artificiale: principi, diritti e raccomandazioni*, in *BioLaw Journal – Rivista di BioDiritto*, 3, 2019, 183 ss. Fra i molti altri, cfr. L. FLORIDI ET AL., *4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*, in *Mind and Machine*, (28)4, 2018, 689 ss., in L. FLORIDI, J. COWLS, *A Unified Framework of Five Principles for AI in Society*, in *Harvard Data Science Review*, (1)1, 2019, si aggiunge ai quattro tradizionali principi della bioetica princialista (*beneficence, non-maleficence, autonomy, and justice*) il concetto di *explicability*.