

Intelligenza artificiale, diritti e doveri nella Costituzione italiana

di Luca Rinaldi

Abstract: *Artificial intelligence, rights and duties in the Italian Constitution* –The paper investigates the potential impact of artificial intelligence (AI) on constitutional rights and duties. Paragraphs II and III briefly analyze the main challenges posed by artificial intelligence for the rights included in the Fundamental Principles and in the First Part of the Italian Constitution. Paragraphs IV, V, VI and VII theorize three new rights whose guarantee appears desirable in order to address the new risks raised by artificial intelligence: the right to an AI disclosure; the right to an explanation; the right to a human in the loop. Their hypothetical legal recognition by EU public law is investigated, also regarding the EC Proposal of an Artificial Intelligence Act. Lastly, paragraph VIII deals with the evolution of constitutional duties caused by the development of artificial intelligence.

Keywords: Artificial Intelligence, Fundamental Rights; Constitutional duties; New Rights; Artificial Intelligence Act.

201

1. Introduzione

Il presente contributo analizza i possibili effetti dell'avvento dell'intelligenza artificiale in materia di diritti e doveri costituzionali.

La prima parte del lavoro (paragrafi II e III) prende sinteticamente in esame alcune delle questioni più urgenti sollevate dalla diffusione di tecnologie basate sull'intelligenza artificiale per i diritti riconosciuti dai Principi Fondamentali e dalla Parte I della Costituzione italiana. La seconda parte (paragrafi III, IV, V e VI) invece, teorizza il riconoscimento di diritti fondamentali ulteriori a quelli esplicitati nella Carta, la cui tutela appare necessaria alla luce della rivoluzione tecnologica. In particolare, sono analizzati tre ipotetici *nuovi diritti*, prendendone in considerazione scenari applicativi e necessità di bilanciamento con altre prerogative: il diritto a conoscere la natura artificiale di un sistema; il diritto alla spiegazione degli output degli algoritmi; il diritto al controllo umano sulle tecnologie intelligenti. È indagata, inoltre, la possibilità di una loro positivizzazione, facendo riferimento principalmente al diritto europeo e alla recente Proposta di Regolamento della Commissione in materia di intelligenza artificiale. L'ultima parte (paragrafo VIII) analizza le nuove declinazioni dei doveri costituzionali connesse all'intelligenza artificiale.

2. Intelligenza artificiale, principio lavorista e diritto al lavoro

Il dibattito riguardante gli effetti sul mercato del lavoro dell'automazione – intesa nel senso lato di partecipazione delle macchine ai processi produttivi – è vecchio quasi quanto la società industriale. Pressoché ogni innovazione tecnologica, infatti, è stata accompagnata dal timore di una diminuzione della forza lavoro, compreso, nel passato recente, il personal computer¹. È necessario premettere un dato storico: tali preoccupazioni si sono quasi sempre rivelate infondate, almeno se si assume come oggetto di analisi il mercato del lavoro nel suo complesso. L'innovazione tecnologica, anche quando ha causato trasformazioni radicali del tessuto produttivo, ha portato alla parallela trasformazione di tempi, spazi e mansioni di lavoro, e non a una diminuzione del numero assoluto degli occupati².

In via di estrema semplificazione, gli studi che tentano di prevedere l'impatto dell'intelligenza artificiale sul mercato del lavoro delle società avanzate si dividono in tre filoni. In primo luogo, vi è chi ipotizza che, analogamente a quanto avvenuto in passato, l'intelligenza artificiale porterà a un cambiamento della domanda di lavoro e delle abilità richieste agli occupati, senza provocarne la diminuzione complessiva³. In secondo luogo, vi è chi vede nell'intelligenza artificiale un'evoluzione tecnologica qualitativamente diversa dalle precedenti, per la capacità di automatizzare molte attività umane intellettuali e creative⁴. Al contrario che in passato, allora, l'intelligenza artificiale sarebbe destinata a privare molti lavoratori delle loro mansioni senza destinarli ad altre, causando spinose questioni sociali. Infine, vi è chi sostiene che l'intelligenza artificiale porterà a una positiva diminuzione dei tempi di lavoro e alla liberazione dalle mansioni più faticose e ripetitive⁵. Secondo questa impostazione, la sua diffusione genererà una società più florida ed equa, trasformando l'attuale necessità di produrre benessere nell'esigenza di redistribuire quello, sovrabbondante, prodotto dalle macchine.

Deve notarsi che tutte queste ipotesi prevedono l'esclusione dal novero degli occupati di diversi di coloro che attualmente vi appartengono. Molti lavoratori a bassa e media specializzazione, infatti, sono destinati ad essere, in ogni caso, espulsi dai processi produttivi, a prescindere dalla loro

¹ Si veda l'analisi del *Time* su timori e speranze che accompagnarono l'avvento del personal computer, risalente al 1965, e l'eloquente commento pubblicato quarant'anni dopo, v. *The Cybernated generation*, in *Time*, 14, 1965, 84-87 e L. Rothman, *This 50-year-old prediction about computers will make you sad*, in *Time.com*, 2 aprile 2015, in time.com/3754781/1965-predictions-computers/.

² D. H. Autor, *Why are there still so many jobs? The history and future of workplace automation*, in *Journal of Economic Perspectives*, 29, 2015, 3 ss; cfr. anche il contributo, qui pubblicato, di M. Borzaga.

³ Cfr. ancora D. H. Autor, *Why are there still so many jobs?*, cit.

⁴ Cfr. M. Ford, *Rise of the Robots: Technology and the Threat of a Jobless Future*, New York, 2015; D. M. West, *What happens if robots take the jobs? The impact of emerging technologies on employment and public policy*, in *Center of Technology Innovation at Brookings*, ottobre 2015, in bit.ly/3Gue9yS.

⁵ B. Arthur, *The second economy*, in *McKinsey Quarterly*, ottobre 2011, mck.co/3BodpaQ.

eventuale sostituzione con altri – più giovani, formati e versatili – impiegati in mansioni diverse. Dal punto di vista della nostra tradizione costituzionale, che pone il lavoro al centro dello sviluppo della persona⁶, ciò non può essere privo di conseguenze. Com'è noto, dal diritto-dovere al lavoro, come declinato all'art. 4 della Costituzione, discende l'obbligazione per la Repubblica di creare condizioni favorevoli per l'occupazione e l'inserimento del maggior numero possibile di cittadini nel tessuto produttivo⁷. Sarà compito delle istituzioni, allora, predisporre un sistema di formazione costante dei lavoratori che permetta il reimpiego di coloro che dovessero perdere l'occupazione a causa della rivoluzione tecnologica. Parallelamente, sarà necessario lo sviluppo di imponenti strutture di welfare redistributivo, al fine di generare una società in cui la piena occupazione non rappresenti più un fine tendenziale dell'ordinamento. A prescindere da quali saranno gli effetti finali in termini macroeconomici, infatti, è evidente che l'intelligenza artificiale è destinata ad innescare processi trasformativi imponenti, di fronte ai quali il mercato del lavoro troverà un nuovo punto d'equilibrio in termini non brevi. Solo se la Repubblica saprà accompagnarvi misure altrettanto incisive il mandato costituzionale potrà dirsi rispettato⁸.

3. L'intelligenza artificiale e il catalogo dei diritti nella Parte I della Costituzione: cenni alle sfide principali

203

Come già evidenziato, l'intelligenza artificiale porrà, nei prossimi anni, questioni inedite in molti ambiti della società, destinate a tradursi in sfide

⁶ Si rimanda innanzitutto a C. Mortati, *Commento all'art. 1 della Costituzione* e G. F. Mancini, *Commento all'art. 4*, in G. Branca (a cura di), *Commentario della Costituzione*, vol. I, art. 1-12: *principi fondamentali*, Bologna, 1975, 1 ss. e 199 ss. *Ex multis*, v. ancora C. Mortati, *Il diritto al lavoro secondo la Costituzione della Repubblica*, in *Raccolta di scritti*, 3, Milano, 1972, 144 ss.; oltre a V. Crisafulli, *Appunti preliminari sul diritto al lavoro nella Costituzione*, in *Rivista giuridica del lavoro*, I, 1951, 161 ss.; R. Scognamiglio, *Lavoro – disciplina costituzionale*, in *Enciclopedia Giuridica Treccani*, XVIII, 1990, 1 ss.; L. Paladin, *I "principi fondamentali" della Costituzione repubblicana: una prospettiva storica*, in *Giur. cost.*, 5, 1997, 3029 ss.; M. Luciani, *Radici e conseguenze della scelta costituzionale di fondare la Repubblica democratica sul lavoro*, in *ADL*, 3, 2010, 629 e ss.; C. Cerutti, *Il lavoro come fondamento della Repubblica democratica italiana*, in *Il lavoro nella giurisprudenza*, 11, 2012, 1061 ss.

⁷ È pacifico che la Costituzione rimetta alla discrezionalità del Legislatore tempi e modi di attuazione del diritto al lavoro, sottoponendone le scelte al controllo di ragionevolezza (v. ad esempio Corte cost. 33/1960; Corte cost. 303/2011; Corte cost. 166/2012) e che la garanzia costituzionale del diritto al lavoro non sia, in via generale, azionabile in giudizio dal singolo cittadino privo d'occupazione. Cfr., oltre ai contributi citati sub. n. 4, L. Tria, *La giurisprudenza della Corte Costituzionale in materia di tutela del diritto al lavoro. Linee di tendenza*, in *Corte costituzionale – studi e documentazione*, 2015, in [bit.ly/3nLCckf](https://doi.org/10.1017/3nLCckf).

⁸ Svolge considerazioni simili C. Casonato, *Costituzione e intelligenza artificiale: un'agenda per il prossimo futuro*, in *BioLaw Journal – Rivista di BioDiritto*, Special Issue 2, 2019, 4-5. V. anche, in generale, L. Tria, *Intelligenza artificiale e diritto al lavoro*, in *I diritti dell'uomo*, 2020, 2, 389 ss.

altrettanto nuove per la tutela dei diritti costituzionali.

Prendendo in considerazione, ad esempio, i diritti attinenti alla sfera della libertà individuale, è evidente come lo sviluppo di tecnologie intelligenti finalizzate al controllo e alla sorveglianza non sia privo di conseguenze per la libertà personale, la libertà di circolazione, la segretezza della corrispondenza. Infatti, strumenti di tracciamento, riconoscimento facciale⁹ e c.d. *predictive policing*¹⁰ sempre più raffinati paiono in grado, da un lato, di garantire un'efficacia senza precedenti alla repressione criminale, dall'altro, di generare meccanismi di controllo diffuso sulla generalità dei cittadini, a prescindere da concreti requisiti di pericolo¹¹. Né si può tacere che, in altri ordinamenti, alcune di queste tecnologie hanno finito per produrre risultati discriminatori, replicando i pregiudizi contenuti nei *dataset* di addestramento e attribuendo, ad esempio, rischi di recidiva maggiori a imputati di colore a parità di storia criminale¹².

Nell'ambito dei rapporti civili, l'intelligenza artificiale fa sorgere nuovi problemi per la garanzia del diritto alla libera manifestazione del pensiero¹³. Social network e provider di servizi mail e messaggistica, infatti, fanno un uso crescente di tecnologie basate sull'elaborazione del linguaggio naturale per l'analisi di contenuti e comunicazioni dei loro utenti. Ciò porta, da un lato, a un'inedita compressione del menzionato diritto alla segretezza della corrispondenza, dall'altro, all'automazione – e alla concentrazione in poche e potenti mani private¹⁴ – di un'attività sostanzialmente censoria come la *content moderation* dei *social network*, peraltro con un margine d'errore, in taluni casi, estremamente ampio¹⁵.

Le possibilità di profilazione dell'individuo dischiuse dall'intelligenza

⁹ Cfr. S. Caines, *The many faces of facial recognition*, in R. Vogl (a cura di), *Research Handbook on Big Data Law*, Cheltenham Glos, 2021, 29 ss.

¹⁰ A. Meijer, M. Wessels, *Predictive policing: review of benefits and drawbacks*, in *International Journal of Public Administration*, 12, 2019, 1031 ss.; W. Perry et al., *Predictive policing. The role of crime forecasting in law enforcement operations*, Washington, 2013.

¹¹ V. ad es. S. Quattrocchio, *Equo processo penale e sfide della società algoritmica*, in *Rivista di BioDiritto – BioLaw Journal*, 1, 2019, 135 ss. e Ead., *Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs. rischi e paure della giustizia digitale "predittiva"*, in *Cassazione penale*, 4, 2019, 1748 ss. Cfr. anche il contributo, pubblicato in questo fascicolo, di E. Carpanelli.

¹² È l'ormai celebre caso del software predittivo COMPAS, che, nel valutare le probabilità di recidiva, replicava i pregiudizi razziali insiti nel sistema giudiziario statunitense, v. *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), in *Harvard Law Review*, 130, 2017, 1530 ss.

¹³ *Ex multis* v. G. Pitruzzella, O. Pollicino, S. Quintarelli, *Parole e potere. Libertà d'espressione, hate speech e fake news*, Milano, 2017; cfr. anche il contributo, pubblicato in questo fascicolo, di M. Tomasi e C. M. Reale.

¹⁴ Cfr. *ex multis* D. Keller, *Making Google the censor*, in *The New York Times*, 12 giugno 2017; M. Monti, *Privatizzazione della censura e internet platforms: la libertà di espressione e i nuovi censori dell'agorà digitale*, in *Rivista italiana di informatica e diritto*, 1, 2019.

¹⁵ E. Llansò et al., *Artificial intelligence, content moderation and freedom of expression*, Working Paper – *Transatlantic Working Group on content moderation online and freedom of expression*, 2019.

artificiale pongono rischi anche per l'effettiva garanzia dei diritti afferenti ai rapporti politici. Non è forse lecito chiedersi, infatti, se le tecniche di manipolazione rese possibili dalla certosa personalizzazione dei contenuti cui ciascun utente è esposto nelle reti sociali abbiano conseguenze sullo stesso genuino esercizio del diritto di voto, o sulla scelta di aderire a partiti e associazioni politiche¹⁶? I recenti fatti di Capitol Hill sono solo un esempio di come la polarizzazione politica nei *social network*, generata attraverso la dolosa sovraesposizione del cittadino a contenuti ideologicamente orientati, possa avere effetti devastanti nel mondo reale¹⁷.

Per quanto riguarda i diritti sociali, sarà necessario, in molti casi, adattarne le modalità di garanzia, al fine di cogliere le opportunità connesse all'innovazione tecnologica e mitigare al contempo i rischi. Le applicazioni dell'intelligenza artificiale, infatti, sono destinate ad investire pressoché ogni settore dell'economia. Il suo utilizzo in ambito medico, ad esempio, se opportunamente regolato, permetterà di incrementare qualità e quantità delle prestazioni erogate dal sistema sanitario nazionale e aumentare la garanzia in concreto del diritto alla salute¹⁸. Allo stesso modo, gli obiettivi previsti dalla Carta fondamentale in materia di rapporti economici, come il razionale sfruttamento del suolo o la tutela del credito, dovranno confrontarsi con le opportunità e le sfide poste dall'impiego dell'intelligenza artificiale in ambito agricolo o nel sistema bancario¹⁹. Lo stesso può dirsi, limitandoci ad un ultimo esempio, del diritto all'istruzione e dell'obbligo in capo alla Repubblica di garantirne l'effettività. L'impiego dell'intelligenza artificiale in scuole e università andrà opportunamente regolato e i programmi formativi adeguati all'innovazione che stiamo vivendo, al fine di diffondere nella società le competenze che in futuro

¹⁶ Il fenomeno è spesso indicato con l'espressione *bubble democracy*. In generale, v. E. Parisier, *The filter bubble: what the internet is hiding from you*, New York, 2011; C. Sunstein, *#republic: Divided democracy in the age of social media*, Princeton, 2017; D. Palano, *La bolla mortale della nuova democrazia*, in *il Foglio*, 1 maggio 2017; M. Fasan, *Intelligenza artificiale e pluralismo: uso delle tecniche di profilazione nello spazio pubblico democratico*, in *BioLaw Journal – Rivista di BioDiritto*, 1, 2019; C. Casonato, *Costituzione e intelligenza artificiale cit.*: p. 713 ss. Cfr., inoltre, il contributo pubblicato in questo fascicolo di L. Sciannella.

¹⁷ Cfr. *U.S. Capitol riot*, in *The New York Times (online)*, [nyti.ms/2TL1iEY](https://www.nytimes.com/2021/01/06/us/politics/capitol-riot.html).

¹⁸ In termini generali cfr. E. Topol, *Deep Medicine. How artificial intelligence can make healthcare human again*, New York, 2019; J. Fernández García et al., *Transforming Healthcare with AI. The impact on workforce and organizations*, Report- EIT Health – Mc Kinsey & C., 2020; Sul legame tra diritto alla salute e organizzazione dei sistemi sanitari, v. L. Busatta, *La salute sostenibile. La complessa determinazione del diritto ad accedere alle prestazioni sanitarie*, Torino, 2018. Cfr. anche S. Amato, *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie*, Torino, 2020 e il contributo, in questo fascicolo, di L. Scaffardi.

¹⁹ Cfr. M. J. Smith, *Getting value from artificial intelligence in agriculture*, in *Animal Production Science*, 60, 2019, 46-54; O. Kaya, *Artificial intelligence in banking*, in *Deutsche Bank Research – EU monitor global financial market*, Francoforte, 2019. Riguardo all'impatto dell'intelligenza artificiale su mercati e finanza si veda anche il contributo, in questo fascicolo, di E. Mostacci.

saranno necessarie²⁰.

4. Intelligenza artificiale e nuovi diritti

Oltre ad indagare l'impatto dell'intelligenza artificiale sul catalogo dei diritti riconosciuti dalla Costituzione, è doveroso chiedersi se l'attuale quadro giuridico sia sufficiente a regolare un'innovazione tecnologica di questa portata.

Il dibattito connesso alla possibilità di considerare tra i diritti inviolabili garantiti dalla Carta anche situazioni giuridiche da essa non esplicitamente previste è fin troppo noto e risalente. Limitandoci a richiamarlo nelle linee essenziali, la dottrina si è a lungo divisa tra chi considerava tassativa l'enumerazione operata dai Costituenti e chi, vedendo nella formula dell'articolo 2 una "clausola aperta", riteneva possibile aggiornare la lista dei diritti²¹. La Corte costituzionale si è dimostrata restia ad aderire a quest'ultima tesi, giungendo, comunque, ad affermarne implicitamente la validità almeno a partire dalla sentenza 561/1987, in cui ha riconosciuto tra i diritti fondamentali quello al controllo sulla sfera sessuale, argomentandone la diretta derivazione dall'art. 2. Deve precisarsi, in ogni caso, che le due impostazioni culturali, per quanto distanti, portano a risultati pratici quasi identici: anche i teorici della "fattispecie chiusa" hanno sempre riconosciuto tutela alle nuove situazioni giuridiche sorte con l'evoluzione della società, riconducendole all'alveo interpretativo di diritti costituzionali già esistenti²².

Deve tenersi a mente quest'ultima circostanza anche nell'affrontare il problema dell'adeguatezza del catalogo dei diritti fondamentali alle sfide poste dall'intelligenza artificiale. Le trasformazioni ad essa connesse, infatti, portano alla necessità di elaborare nuove posizioni giuridiche a protezione dell'individuo. Se si tratti di nuovi diritti fondamentali *tout-court* o estrinsecazioni di diritti già riconosciuti dalla Costituzione è una questione di limitato rilievo pratico. In ambo i casi, ad essere concretamente esistente è la necessità di elaborare, di fronte a nuovi problemi, nuove strategie di tutela. Da qui in poi, l'espressione *nuovi diritti*

²⁰ Cfr. L. Chen, P. Chen, Z. Lin, *Artificial Intelligence in Education: A Review*, in *IEEE Access*, 8, 2020, 75264 ss.

²¹ Possono citarsi, tra i principali teorici della fattispecie aperta, C. Mortati, *La Corte costituzionale e i presupposti della sua vitalità*, in *Iustitia*, 1949, 69 ss., con accezione neogiusnaturalistica, e A. Barbera, *Art. 2*, in G. Branca (a cura di), *Commentario della Costituzione*, Bologna-Roma, 1975, 50 ss., con un'interpretazione che valorizza l'evoluzione dei costumi. Per le possibili incoerenze della teoria dell'apertura, v. A. Pace, *Problematica delle libertà costituzionali. Parte generale*, 1990, 4 ss. e P. Caretti, *I diritti fondamentali. Libertà e Diritti sociali*, Torino, 2005, 137 ss. Sul tema cfr. anche P. Barile, *Diritti dell'uomo e libertà fondamentali*, Bologna, 1984, 56 ss.; M. Cartabia, *I "nuovi" diritti*, in *Le confessioni religiose nel diritto dell'Unione Europea*, Bologna, 2012, 97 ss.; F. Modugno, *I "nuovi diritti" nella giurisprudenza costituzionale*, Torino, 1994, 5 ss.; R. Bin, *Critica della teoria dei diritti*, Milano, 2018, p. 55 ss.

²² Sul punto v. ancora A. Pace, *Problematica*, cit. e P. Caretti, *I diritti*, cit.

sarà utilizzata tenendo a mente queste premesse, senza la volontà di dare risposta definitiva a una questione che supera di gran lunga portata e scopi dello scritto.

Tanto premesso, l'intelligenza artificiale pone tre problemi essenziali cui pare opportuno rispondere con categorie giuridiche *ad hoc*. In primo luogo, la sua stessa riconoscibilità, che porta a ipotizzare un diritto ad essere informati dell'utilizzo di tecnologie intelligenti. In secondo luogo, la trasparenza dei sistemi, intesa come possibilità di interpretarne gli *output*, che porta a teorizzare un diritto alla spiegazione di quest'ultimi. Infine, la necessità di conservare nelle mani dell'uomo il governo delle tecnologie intelligenti, da cui deriverebbe, appunto, un diritto al controllo umano su di esse²³. Peraltro, si tratta di prospettazioni giuridiche inizialmente teorizzate in dottrina ed enunciate in documenti di soft-law, ma che stanno trovando riconoscimento, in varie forme, anche in atti normativi vincolanti, specialmente a livello europeo²⁴.

5. Il diritto ad essere informati dell'utilizzo di tecnologie intelligenti

Il tema della distinguibilità dall'essere umano dell'intelligenza artificiale ha interessato gli esperti fin dagli albori della disciplina. Lo stesso Alan Turing, nel teorizzare il noto test per definire se una macchina possa dirsi *pensante*, pose la linea di confine nella capacità di comunicare con degli esseri umani senza venire riconosciuta²⁵.

Si deve evidenziare che nascondere la natura artificiale di un sistema ha effetti non irrilevanti sui risultati dell'interazione con l'umano. Già nel 1970, lo studioso di robotica Masahiro Mori identificò il fenomeno della c.d. *uncanny valley*: la quota di empatia e gradimento verso un robot antropomorfo cresce proporzionalmente alla sua verosimiglianza, ma cala bruscamente quando questa raggiunga livelli tali da confondere l'utente²⁶.

²³ Si tratta di garanzie teorizzate anche da C. Casonato, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo*, numero speciale, 2019, 125 ss. e Id., *Costituzione e intelligenza artificiale*, cit., 722 ss. Si vedano, inoltre, le considerazioni dello stesso autore nell'introduzione di questo fascicolo.

²⁴ Il riferimento, nelle forme che si diranno, è al Reg. UE 679 del 2016 (c.d. GDPR) e alla recente Proposta di Regolamento sull'intelligenza artificiale presentata il 21 aprile del 2021 dalla Commissione Europea (COM(2021) 206 final). Per quanto riguarda le fonti di soft-law, invece, vengono principalmente in rilievo gli *Orientamenti etici per un'IA affidabile* elaborati dal Gruppo di Esperti d'Alto Livello sull'IA dalla Commissione Europea e resi pubblici l'8 aprile 2019. Sul piano comparato, possono segnalarsi il *Bolstering Online Transparency Act*, emanato nel 2018 dallo stato americano della California, in materia di *disclosure* dell'utilizzo di bot online, e la *Directive on Automated Decision-Making* entrata in vigore in Canada nell'aprile 2019, che prevede specifici requisiti di spiegabilità e controllo umano.

²⁵ A. Turing, *Computing machinery and intelligence*, in *Mind*, 236, 1950, 433 ss.

²⁶ M. Mori, *The uncanny valley*, in *IEEE Robotics & Automation Magazine*, 2, 2012, 98 ss. (originale *Bukimi no tani*, in *Energy*, 4, 1970, 33 ss.).

Inoltre, non esplicitare l'uso di tecnologie intelligenti parrebbe portare a vantaggi commerciali: stando ad uno studio del 2019, l'uso di chatbot nella vendita a distanza porta a risultati assimilabili a quelli ottenuti da venditori esperti. Se, però, la natura artificiale dell'interlocutore è rivelata al principio della conversazione, l'efficacia del chatbot cala drasticamente, portando a una diminuzione delle vendite del 79%²⁷. La *disclosure* dell'utilizzo di intelligenza artificiale, dunque, potrebbe apparire sconveniente a chi agisca per finalità commerciali o, più genericamente, di persuasione. Deve aggiungersi lo stato ancora acerbo delle tecnologie in grado di mimetizzare la loro natura artificiale: i risultati più importanti, finora, sono stati raggiunti proprio in materia di chatbot e assistenti vocali²⁸. I problemi ad esse connessi, prevedibilmente, cresceranno di pari passo al loro sviluppo. Ciò nonostante, non può negarsi che scenari inquietanti si siano già realizzati: si pensi al massiccio utilizzo di *bot* sui social media al fine di inquinare il dibattito politico con finti messaggi e *fake news* o alla diffusione di c.d. *deep fake*, anche pornografici, ritraenti persone reali in situazioni mai avvenute²⁹. Di fronte a questi rischi, e agli altri che potenzialmente potrebbero sorgere in futuro, la previsione di un obbligo generalizzato di dichiarare l'impiego di tecnologie intelligenti potrebbe rappresentare un esempio di regolazione lungimirante e in grado di orientare il progresso tecnologico verso la giusta direzione. In capo all'utente, così, si configurerebbe un diritto ad essere informato della presenza dell'intelligenza artificiale in ogni situazione in cui sia ragionevole ipotizzare che non se ne avvedrebbe altrimenti, eventualmente azionabile in giudizio in caso di violazione³⁰.

Peraltro, a livello europeo la necessità di porre chiari obblighi di *disclosure* dell'intelligenza artificiale è già stata presa in considerazione dal Legislatore. Gli articoli 13.2 lett. f), 14.2 lett. g) e 15.1 lett. h) del GDPR, che disciplinano gli obblighi informativi nei confronti dell'interessato del trattamento a seconda del contesto in cui i dati sono raccolti, impongono di fornire informazioni «sull'esistenza di un processo decisionale

²⁷ X. Luo et al., *Machines vs humans: the impact of artificial intelligence chatbot disclosure on customer purchases*, in *Marketing science*, 6, 937 ss.

²⁸ Non è ancora possibile parlare di indistinguibilità dall'essere umano nei termini teorizzati da Turing, ma la riconoscibilità di chatbot e assistenti vocali spesso richiede uno sforzo irragionevole rapportato al contesto di riferimento, cfr. ad es. D. Adiwardana et al., *Towards a Human-like Open-Domain Chatbot*, 2020, [arXiv:2001.09977](https://arxiv.org/abs/2001.09977); J. P. McIntire, L. K. McIntire, P. R. Havig, *Methods for chatbot detection in distributed text-based communications*, *International Symposium on Collaborative Technologies and Systems*, 2010, 463 ss.

²⁹ D. K. Citron, R. Chesney, *Deep fakes: a looming challenge for privacy, democracy, and national security*, in *California Law Review*, 107, 2019, 1753 ss.

³⁰ Cfr. ancora C. Casonato, *Intelligenza artificiale e diritto costituzionale*, cit., 125 ss. e Id., *Costituzione e intelligenza artificiale*, cit., 722 ss.; oltre al contributo introduttivo di questo fascicolo.

automatizzato» che coinvolga i dati trattati³¹. Non può parlarsi, in ogni caso, di un pieno riconoscimento del diritto ad essere informati dell'impiego di tecnologie intelligenti. Tali norme, infatti, impongono di rivelare l'utilizzo di intelligenza artificiale solo nel caso di decisioni automatizzate: un ambito di applicazione limitato e insufficiente ad affrontare la varietà e la complessità dei problemi che la mancata *disclosure* può generare³². La Proposta di Regolamento sull'intelligenza artificiale presentata dalla Commissione Europea nell'aprile 2021³³ pare, però, farsi carico del problema. L'art. 52, infatti, impone di informare l'utente della presenza di un sistema intelligente, qualora non sia evidente, ogni volta che questo sia diretto all'interfaccia con persone fisiche; abbia finalità di riconoscimento delle emozioni o identificazione e catalogazione biometrica; sia stato utilizzato per generare *deepfake*³⁴. Il quadro normativo europeo, allora, sembra andare verso il pieno riconoscimento del diritto in esame e non resta che attendere il decorso dell'iter di applicazione della norma.

6. Trasparenza algoritmica e diritto alla spiegazione

La storia dell'intelligenza artificiale è stata segnata, inizialmente, dallo sviluppo di c.d. sistemi basati sulla conoscenza, in grado di svolgere compiti associati all'intelligenza umana grazie alla programmazione delle

³¹ Tutte e tre le norme impongono di fornire all'interessato informazioni sulla «esistenza di un processo decisionale automatizzato, compresa la profilazione di cui all'articolo 22, paragrafi 1 e 4, e, almeno in tali casi, informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato».

³² Ad esempio, l'utilizzo di bot con finalità distorsive del dibattito politico, o la generazione di contenuti *deep fake*, non paiono riconducibili, normalmente, all'ambito applicativo della «decisione automatizzata».

³³ Per dei primi commenti si vedano C. Casonato, B. Marchetti, *Prime osservazioni sulla Proposta di Regolamento dell'Unione Europea in materia di intelligenza artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, 3, 2021; N. A. Smuha et al., *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act*, 2021, in [dx.doi.org/10.2139/ssrn.3899991](https://doi.org/10.2139/ssrn.3899991); P. Hacker, *A legal framework for AI training data – from first principles to the Artificial Intelligence Act*, 2021, [doi:10.1080/17579961.2021.1977219](https://doi.org/10.1080/17579961.2021.1977219); A. Lavorgna, G. Suffia, *La nuova proposta europea per regolamentare i Sistemi di Intelligenza Artificiale e la sua rilevanza nell'ambito della giustizia penale: un passo necessario, ma non sufficiente, nella giusta direzione*, in *Diritto Penale Contemporaneo*, 2, 2021, 88 ss.; F. C. La Vattiatà, *Brevi note "a caldo" sulla recente proposta di Regolamento UE in tema di intelligenza artificiale*, in *Diritto Penale e Uomo*, 6, 2021. Sull'approccio europeo alla regolazione dell'intelligenza artificiale si veda anche il contributo, in questo fascicolo, di B. Marchetti e L. Parona.

³⁴ La norma, bilanciando il diritto in esame con altri interessi considerati di pari rango, specifica che la *disclosure* non è obbligatoria qualora il sistema sia autorizzato dalla legge e destinato ad accertare, prevenire e indagare reati, o, nel caso di *deepfake*, quando la mancata rivelazione costituisca esercizio del diritto alla libertà di espressione e del diritto alla libertà delle arti e delle scienze, fatte salve le tutele adeguate per i diritti e le libertà dei terzi.

regole e nozioni necessarie³⁵. Il principale limite di questo approccio consisteva nella necessità di sviluppare, anche per il raggiungimento di risultati molto semplici, programmi estremamente complessi. Ciò nonostante, è da questo filone che derivano i c.d. sistemi esperti, programmi operanti in campi circoscritti e specifici che, a cavallo tra gli anni '70 e '80, rappresentarono le prime applicazioni dell'IA ad ottenere effettivo impiego industriale³⁶.

Fin dagli esordi della disciplina, la ricerca nel campo dell'intelligenza artificiale tentò anche altre strade. Una di queste mirava allo sviluppo di c.d. reti neurali artificiali, muovendo dal presupposto che la mimesi della struttura essenziale del cervello avrebbe permesso di sviluppare programmi in grado di apprendere e perfezionare il loro funzionamento con l'esperienza³⁷. Tale filone di ricerca fu presto abbandonato a causa della quantità di dati necessaria all'addestramento delle reti, irrealistica per le potenzialità dei computer degli anni '60, '70 e '80³⁸. Più recentemente, però, l'aumento della capacità computazionale degli elaboratori ha portato al vertiginoso sviluppo del settore dell'apprendimento automatico, dimostratosi particolarmente efficace negli ambiti in cui i sistemi basati sulla conoscenza avevano mostrato i limiti maggiori³⁹. A partire dagli anni '10 del duemila, inoltre, si è affermata una tipologia di reti neurali particolarmente complessa ed efficace, il c.d. *deep learning*⁴⁰.

Se i sistemi esperti replicavano l'applicazione delle conoscenze specialistiche relative a un determinato settore, lo stesso non può dirsi delle reti neurali. Il legame con la conoscenza è, in questo caso, molto più debole: il funzionamento delle reti si basa sull'osservazione dei dati d'addestramento. A determinare l'output di un sistema basato sul *deep learning* è l'analisi di correlazioni statistiche in un database la cui misura supera di diversi ordini di grandezza le capacità della mente umana. Ne deriva l'impossibilità di ricavare una spiegazione riconducibile all'applicazione di conoscenze teoriche o pratiche, caratteristica cui è comune riferirsi con l'espressione inglese *black-box*⁴¹. Devono aggiungersi

³⁵ Cfr. *ex multis* S. J. Russell, P. Norvig, *Artificial intelligence: a modern approach* (4^a ed.), Hoboken (NJ), 2020, 17 ss. In generale, v. anche il contributo introduttivo pubblicato *supra* di P. Traverso.

³⁶ N. J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge, 2009, 229 ss.

³⁷ Gli studi più risalenti nell'ambito risalgono addirittura agli anni '40, cfr. W. S. McCulloch, W. Pitts, *A logical calculus of ideas immanent in nervous activity*, in *Bulletin of Mathematical Biophysics*, 5, 1943, 115 ss.

³⁸ S.J. Russell, P. Norvig, *Artificial Intelligence*, cit., 19 ss.

³⁹ È il caso, ad es., del riconoscimento di immagini o della traduzione automatica, v. A. Krizhevsky, I. Sutskever, G. E. Hinton, *ImageNet classification with deep convolutional neural networks*, in *NIPS*, 1, 2012.

⁴⁰ V. in generale C. C. Aggarwal, *Neural networks and deep learning – a textbook*, Berlino, 2018.

⁴¹ V. F. Pasquale, *The Black-Box Society: The Secret Algorithms That Control Money and Information*, Harvard, 2016; R. V. Yampolskiy, *Unexplainability and incomprehensibility of*

tre precisazioni: in primo luogo, in alcuni settori, come il riconoscimento di immagini, il *deep learning* è in grado di fornire risultati comparabili o superiori a quelli dei migliori esperti umani⁴². In secondo luogo, i risultati delle reti neurali hanno sempre un margine d'errore, anche se spesso infinitesimale, dovuto alla stessa natura statistica del loro funzionamento⁴³. Infine, le reti neurali sono esposte al rischio di replicare, nei loro output, distorsioni presenti nei dati di addestramento o apprese nel corso del loro funzionamento (si pensi ai casi di *bias* discriminatori già menzionati nei paragrafi precedenti)⁴⁴.

L'utilizzo di algoritmi *black-box* appare a molti non accettabile in diverse circostanze, per l'ampiezza della delega alla macchina che ne deriva. Così, se da un lato la ricerca in ambito tecnologico dirige i propri sforzi verso lo sviluppo di una c.d. *explainable artificial intelligence*⁴⁵, dal punto di vista del giuridico pare necessario prescrivere requisiti minimi di trasparenza delle tecnologie intelligenti. A tal proposito, potrebbe proporsi, a tutela di un incompressibile spazio di indipendenza e autonomia dell'essere umano, il riconoscimento di un diritto ad ottenere una spiegazione dei risultati di un sistema⁴⁶.

Del resto, si tratta di un'esigenza già avvertita dal regolatore europeo. La trasparenza dei sistemi, infatti, è tra i principi chiavi degli *Orientamenti Etici per un'intelligenza artificiale affidabile* del 2019, elaborati dal Gruppo di Esperti di Alto Livello sull'IA nominato dalla Commissione Europea⁴⁷. Inoltre, i già citati articoli 13.2 lett. f), 14.2 lett. g) e 15.1 lett. h) del GDPR, nello stabilire gli obblighi informativi in capo al titolare del trattamento, impongono che, in caso di processo decisionale automatizzato, all'interessato siano fornite «informazioni significative sulla logica utilizzata», con l'intento di garantire un coefficiente almeno parziale di

artificial intelligence, 2019, [arXiv:1907.03869](https://arxiv.org/abs/1907.03869); A. Adadi, M. Berrada, *Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)*, in *IEEE Access*, 6, 2018, 5244-5245.

⁴² Cfr. ad es. A. Hekler et al., *Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images*, in *European Journal of Cancer*, 118, 2019, 91 ss.

⁴³ Si rimanda a S. Quintarelli (a cura di), *Intelligenza artificiale: cos'è davvero, come funziona, che effetti avrà*, Torino, 2020.

⁴⁴ B. Friedman, H. Nissenbaum, *Bias in computer systems*, in *ACM Transactions on Information Systems*, 3, 1996, doi.org/10.1145/230538.230561; R. Dobbe et al., *A Broader View on Bias in Automated Decision-Making: Reflecting on Epistemology and Dynamics*, 2018, [arXiv:1807.00553](https://arxiv.org/abs/1807.00553).

⁴⁵ Cfr. G. Vilone, E. Longo, *Explainable artificial intelligence: a systematic review*, 2020, [arXiv:2006.00093](https://arxiv.org/abs/2006.00093).

⁴⁶ Così C. Casonato, *Intelligenza artificiale e diritto costituzionale*, cit., 125 ss. e Id., *Costituzione e intelligenza artificiale*, cit., 722 ss. Sul tema cfr. anche, da un diverso punto di vista, le considerazioni in materia di principio di trasparenza di M. Fasan, nel contributo pubblicato in questo fascicolo, e le osservazioni dello stesso C. Casonato, nell'introduzione.

⁴⁷ Cfr. p. 20 ss.

trasparenza⁴⁸. Ad ogni modo, anche questo onere è previsto unicamente nel ristretto ambito applicativo dei trattamenti di dati personali finalizzati alla decisione automatizzata⁴⁹. Non può dirsi, dunque, che il GDPR garantisca un completo diritto alla spiegazione.

Preme evidenziare che, allo stato dell'arte, la scelta di optare per un sistema trasparente in luogo di uno totalmente o parzialmente opaco non è priva di conseguenze: in diversi settori esiste un vero e proprio *trade-off* tra efficacia e trasparenza dell'algoritmo⁵⁰. Il diritto alla spiegazione, dunque, dovrà essere declinato tenendone in considerazione la fattibilità tecnologica e la grande varietà di applicazioni in cui potrebbe venire in gioco. Infatti, pare doveroso garantire una spiegabilità pressoché completa, anche al costo di rinunciare all'utilizzo di tecnologie intelligenti, in ambiti cruciali come la decisione giudiziaria, amministrativa, o in altri settori in cui vengano in gioco rilevanti interessi individuali (si pensi, per limitarsi a un solo esempio, all'accesso al credito)⁵¹. È ragionevole, però, giungere a conclusioni diverse quando dall'opacità dell'algoritmo derivino rischi limitati (può essere il caso di un sistema che supporta l'utente di un motore di ricerca)⁵² o quando la rilevanza degli interessi coinvolti porti a ritenere accettabile un livello di trasparenza minore, pur di trarre ogni profitto dall'efficacia dell'algoritmo (si pensi a taluni sistemi di riconoscimento delle immagini, che permettono diagnosi precoci di gravi patologie)⁵³. Il diritto

⁴⁸ Per il testo delle norme v. *supra*, nota 31. L'ambigua formula adottata dal Legislatore europeo è stata ampiamente commentata in dottrina, v. F. Pizzetti, *La protezione dei dati personali e la sfida dell'intelligenza artificiale*, in Id. (a cura di), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018, 34 ss. e S. Wachter, B. Mittelstadt, L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, in *International Data Privacy Law*, 2, 2017, 76 ss.

⁴⁹ Può farsi l'esempio di algoritmi impiegati nella robotica industriale, difficilmente riconducibili all'ambito della decisione automatizzata sulle persone fisiche, ma la cui eventuale opacità di funzionamento potrebbe risultare problematica.

⁵⁰ R. V. Yampolskiy, *Unexplainability and incomprehensibility*, cit.; A. Adadi, M. Berrada, *Peeking Inside the Black-Box*, cit.; S. Sarkar et al., *Accuracy and Interpretability Trade-offs in Machine Learning Applied to Safer Gambling*, in *CEUR Workshop Proceedings*, 2016, 1773 ss.

⁵¹ Sul tema si vedano le considerazioni di L. Ammannati, G. L. Greco, *Il credit scoring alla prova dell'intelligenza artificiale*, in U. Ruffolo (a cura di), *XXXVI Lezioni di diritto dell'intelligenza artificiale*, Torino, 2021, 373 ss. Inoltre, sul tema dell'utilizzo dell'IA nella giustizia e nella pubblica amministrazione si vedano, rispettivamente, i contributi qui pubblicati di S. Penasa e M. Macchia.

⁵² Cfr. A. Bunt, M. Lount, C. Lauzon, *Are explanations always important? A study of deployed, low-cost intelligent interactive systems*, in *Proceedings of the 2012 ACM international conference on IUI*, 2012, 169 ss.

⁵³ Cfr. i risultati ottenuti da P. Rajpurkar et al., *CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning*, 2017, arXiv:1711.05225 e A. Hekler et al., *Deep learning*, cit. L'impiego dell'intelligenza artificiale nella diagnostica deve, in ogni caso, essere accompagnato da cautele, poste le conseguenze di eventuali malfunzionamenti, le possibili ripercussioni sul rapporto medico-paziente, i rischi di *deskilling* e di adesione acritica al sistema da parte del medico (c.d. distorsione dell'automazione). Cfr. S. Triberti et al., *Is Explanation a Marketing Problem? The Quest*

in esame dovrebbe, in sintesi, sostanzarsi nel diritto alla miglior spiegazione possibile, da bilanciare nel caso concreto con le altre prerogative in gioco, al pari, del resto, di ogni altro diritto fondamentale⁵⁴.

Pare questa la direzione in cui si muove anche la Proposta di Regolamento in materia di intelligenza artificiale della Commissione Europea, il cui art. 13.1 prevede che i sistemi considerati ad alto rischio siano «progettati e sviluppati in modo tale da garantire che il loro funzionamento sia sufficientemente trasparente da consentire agli utenti di interpretare l'output del sistema e utilizzarlo adeguatamente». La particolare formulazione legislativa, che pone l'obiettivo di un sistema *sufficientemente trasparente*, e il riferimento ai soli sistemi ad alto rischio (il cui ambito è, comunque, molto ampio)⁵⁵ paiono idonei a configurare, qualora la norma fosse effettivamente approvata, un diritto alla spiegazione da un lato non limitato a spazi di applicazione troppo angusti e dall'altro in grado, come qui si auspica, di adattarsi alla varietà dei possibili contesti di riferimento.

7. Il diritto al controllo umano sul sistema

L'altra garanzia che pare opportuno elevare al rango di diritto fondamentale è il mantenimento di un determinato livello di controllo in capo a un essere umano sull'attività di un sistema di intelligenza artificiale (c.d. requisito dello *human in the loop* o *human oversight*)⁵⁶. Il tema del controllo umano viene in gioco, come diritto individuale, in primo luogo quando le tecnologie intelligenti siano impiegate in processi decisionali: scenari nei quali il diritto in esame si sostanzia nella garanzia di non essere sottoposto a decisioni prive di un'effettiva componente umana⁵⁷. Questa posizione giuridica ha ottenuto riconoscimento all'art. 22 del GDPR, che al primo paragrafo enuncia il «diritto a non essere sottoposto a una decisione

for Trust in Artificial Intelligence and Two Conflicting Solutions, in *Public Health Genomics*, 23, 2020, 2 ss.

⁵⁴ V. anche le considerazioni di A. Paez, *The Pragmatic Turn in Explainable Artificial Intelligence (XAI)*, in *Minds and Machines*, 29, 2019, 441 ss.

⁵⁵ L'art. 6 della Proposta identifica come ad alto rischio un'ampia serie di applicazioni dell'IA, elencate all'allegato III (es: accesso al credito, gestione dei flussi migratori, identificazione biometrica quando permessa ecc.) e tutti i sistemi di intelligenza artificiale che siano prodotti o componenti di sicurezza di prodotti disciplinati dalla normativa di armonizzazione indicata all'allegato II e per i quali sia prevista una valutazione di conformità da parte di terzi.

⁵⁶ Cfr. M. L. Jones, *The right to a human in the loop: Political constructions of computer automation and personhood*, in *Social Studies of Science*, 47, 2, 2017, 216 ss. Si vedano, inoltre, le osservazioni in materia di sorveglianza umana di M. Fasan, nel contributo pubblicato in questo fascicolo.

⁵⁷ Cfr. C. Casonato, *Costituzione e intelligenza artificiale*, cit., p. 723-724; A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal – Rivista di BioDiritto*, 1, 2019, 79 ss. Inoltre, si veda di nuovo il contributo introduttivo di questo fascicolo dello stesso C. Casonato.

basata unicamente sul trattamento automatizzato»⁵⁸. Deve evidenziarsi, però, che il paragrafo successivo prevede eccezioni estremamente ampie, tali da far dubitare che alla norma residui un effettivo margine di applicazione⁵⁹. La tutela apprestata dall'art. 22, inoltre, appare insufficiente anche dal punto di vista dell'ambito di applicazione: la decisione automatizzata non è l'unico scenario in cui può venire in rilievo il requisito dello *human in the loop*. L'insieme di tecnologie intelligenti con cui l'essere umano si interfaccia è molto più ampio, e pare necessario riconoscere la garanzia in esame anche in altri casi (si pensi, ad esempio, alla robotica industriale in contesti produttivi ibridi)⁶⁰.

Possono individuarsi vari livelli di controllo in capo all'operatore, corrispondenti a un ruolo crescente della componente umana: l'assenza totale di sorveglianza e dominio sul sistema, operante in un contesto totalmente automatizzato (*human out of the loop*); un controllo limitato alla fase di sviluppo e al monitoraggio del funzionamento, con la possibilità di discostarsi ex-post dagli output (*human on the loop*); la possibilità di controllo e intervento in ogni fase del funzionamento del sistema, interrompendone e modificandone il lavoro (*human in the loop* in senso proprio); il controllo pieno del funzionamento del sistema, unito alla possibilità di decidere, caso per caso, tempi e modi dell'utilizzo, in un contesto di automazione estremamente limitata (*human in command*)⁶¹. Inoltre, il requisito del controllo umano è strettamente connesso al livello di trasparenza della tecnologia: i livelli più alti di controllo, che prevedono la possibilità di intervenire sui singoli cicli di funzionamento al fine di modificare gli output, sono possibili solamente se essa sia almeno in parte interpretabile.

Anche nel riconoscimento del diritto allo *human oversight*, allora, dovrà tenersi conto dello stato dell'arte tecnologico, ed esso assumerà un'ampiezza differente a seconda degli interessi posti a bilanciamento. Così, vi saranno contesti in cui apparirà necessario un controllo penetrante sul

⁵⁸ Il paragrafo recita: «L'interessato ha 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». Peraltro, una disposizione analoga era prevista all'art. 15 dell'abrogata direttiva europea 95/46/CE.

⁵⁹ Il secondo paragrafo dell'art. 22, infatti, prevede che il primo non si applichi qualora la decisione sia necessaria all'esecuzione di un contratto tra titolare del trattamento e interessato, sia autorizzata dal diritto dell'Unione o dello stato membro di riferimento o si basi sul consenso dell'interessato. Per un commento v. A. Oddenino, *Decisioni algoritmiche e prospettive internazionali di valorizzazione dell'intervento umano*, in *DPCE online*, 1, 2020, 199 ss; A. Caia, *Art. 22*, in G. M. Riccio, G. Scorza, E. Belisario (a cura di), *GDPR e normativa privacy. Commentario*, 2018, 219 ss.

⁶⁰ Per i potenziali rischi per l'operatore umano cfr. A.F.T. Winfield et al., *Robot Accident Investigation: A Case Study in Responsible Robotics*, in A. Cavalcanti, B. Dongol, R. Hierons, J. Timmis, J. Woodcock (a cura di), *Software Engineering for Robotics*, Cham, 2021, 165 ss.

⁶¹ Una classificazione simile è proposta dagli *Orientamenti etici*, cit., 18.

sistema – e di conseguenza l'impiego delle tecnologie che non lo permettono dovrà escludersi – e altri in cui sarà sufficiente un approccio basato sul monitoraggio esterno, che garantisca solo la possibilità, per l'agente umano, di ignorare o sovvertire l'output⁶². Requisito, quest'ultimo, che si potrebbe considerare il nucleo essenziale del diritto, e che rappresenterebbe un'importante tutela individuale nei contesti in cui l'utilizzo di algoritmi c.d. black-box porti a un sacrificio quasi totale del menzionato diritto alla spiegazione.

Volgendo lo sguardo, anche in questo caso, alle possibilità di una positivizzazione del diritto in esame che superi i visti limiti dell'art. 22 del GDPR, si deve menzionare nuovamente la Proposta di Regolamento della Commissione. L'art. 14 del testo detta requisiti stringenti in materia di sorveglianza e controllo umano sui sistemi di intelligenza artificiale ad alto rischio⁶³. Apprezzabilmente, la norma impone misure che mettano in guardia l'utente dal rischio di fare eccessivo affidamento sul sistema e prescrive che non venga mai meno la possibilità di arrestarne il funzionamento o ribaltarne gli output⁶⁴. Inoltre, alcuni degli accorgimenti previsti intersecano il più volte menzionato requisito tecnico della trasparenza, imponendo, ad esempio, che l'operatore umano sia messo «in grado di interpretare correttamente l'output [...], tenendo conto, in particolare, delle caratteristiche del sistema e degli strumenti e dei metodi di interpretazione disponibili». Da questo punto di vista, nonostante alcune parti della norma invitino a tenere conto dello stato dell'arte tecnologico⁶⁵,

⁶² Si ipotizzano, tra i molti possibili, due esempi attinenti all'arte medica: se il controllo umano pare potersi ridurre al monitoraggio esterno con possibilità di revisione dell'output nei citati esempi di sistemi che coadiuvano la diagnostica per immagini (cfr. n.53) al fine di profittare dei vantaggi connessi alla loro efficienza, deve giungersi a conclusioni diverse nel caso di algoritmi predittivi del decorso clinico dei pazienti volti a razionalizzare l'accesso a trattamenti medici, anche di sostegno vitale. In tal caso, la rilevanza dei diritti coinvolti e i profondi risvolti etici della scelta, in cui al vantaggio di un paziente corrisponde lo svantaggio di un altro, imporrebbero la conservazione, in capo al medico, di un pieno controllo sul sistema e l'esclusione di algoritmi non totalmente spiegabili, non apparendo accettabile il rischio di delegare alla macchina anche una minima componente della decisione. L'attenzione della ricerca per questo tipo di algoritmi è cresciuta negli ultimi anni, in ragione della pandemia di SARS-CoV-2; non risultano, allo stato dell'arte, implementazioni pratiche. Cfr. ad es. D. Patel et al., *Machine learning based predictors for COVID-19 disease severity*, in *Scientific Reports*, 11, 2021, 4673; D. S. Chow et al., *Development and external validation of a prognostic tool for COVID-19 critical disease*, in *PLoS One*, 12, 2020.

⁶³ La norma richiede che l'operatore umano sia messo nelle condizioni di: comprendere capacità e limiti del sistema; monitorarne il funzionamento e percepirne le anomalie; rimanere consapevole del rischio di distorsione dell'automazione; interpretare correttamente gli output, tenendo conto di caratteristiche del sistema e metodi di interpretazione disponibili; decidere di non usare la tecnologia, ignorare o sovvertirne i risultati o interromperne il funzionamento.

⁶⁴ Valorizzano questi elementi C. Casonato, B. Marchetti, *Prime osservazioni sulla Proposta di Regolamento*, cit., 17-18.

⁶⁵ Ad es., oltre al citato riferimento a caratteristiche del sistema e metodi di interpretazione disponibili, la norma impone che le misure da essa prescritte siano

non pare potersi escludere che essa, se rigorosamente interpretata, porti a un eccesso di dettaglio e severità non adatto a tutti gli scenari applicativi, posta l'ampiezza della categoria dei sistemi di intelligenza artificiale considerati ad alto rischio. In caso di approvazione dell'attuale formulazione, allora, sarà la lettura che ne daranno giudici, autorità di settore e operatori pratici a determinare se il regime da essa previsto risulti eccessivamente rigido e difficilmente applicabile o consenta, invece, approcci flessibili come quello auspicato, del resto maggiormente corrispondente alla categoria di diritto fondamentale, in cui, secondo le tesi di questo scritto, la garanzia del controllo umano sul sistema va ricompresa.

8. Intelligenza artificiale e doveri costituzionali

Com'è noto, la Costituzione fa riferimento in vari punti alla categoria dei doveri⁶⁶. Talvolta, per enfatizzare la natura ambivalente di alcune prerogative individuali, al tempo stesso diritto del singolo e dovere verso la comunità: è così per il lavoro, il voto o l'obbligo di mantenere, istruire ed educare i figli. Meno spesso, per stabilire la doverosità di talune prestazioni, a prescindere dal legame immediato con un diritto: è il caso dell'obbligo tributario e dei doveri di difesa della patria e di fedeltà alla Repubblica, con cui si chiude la Parte I della Costituzione. Accanto a queste previsioni, si pone la clausola generale dell'art. 2, che, nel garantire i diritti inviolabili dell'uomo, richiede allo stesso tempo l'adempimento dei «doveri inderogabili di solidarietà politica, economica e sociale».

In letteratura si è evidenziato come, per lungo tempo, valore e portata dei doveri costituzionali siano stati sottovalutati, in ragione del rischio di legittimare, con un'eventuale valorizzazione della categoria, compressioni dei diritti che avrebbero finito per tradire lo spirito della Carta⁶⁷. Questa impostazione non è da condannare totalmente: si pensi alle conseguenze cui avrebbe potuto portare, negli anni dell'emergenza, un'interpretazione estensiva del generico dovere di fedeltà alla

garantite «a seconda delle circostanze».

⁶⁶ In generale v. *ex multis* G. Tarli Barbieri, *Doveri inderogabili*, in S. Cassese (a cura di), *Dizionario di diritto pubblico*, Milano, 2006, 2072 ss.; G. Peces-Barba Martinez, *Diritti e doveri fondamentali*, in *Digesto delle discipline pubblicistiche*, V, Torino, 1990, 159 ss.; A. Cerri, *Doveri pubblici*, in *Enciclopedia giuridica Treccani*, Roma, 1989, XII, 1 ss.; G. M. Lombardi, *Contributo allo studio dei doveri costituzionali*, Milano, 1967; L. Violante, *Il dovere di avere doveri*, Bologna, 2014; A. M. Poggi, *Corte costituzionale e doveri*, in *Corte costituzionale e sistema istituzionale. Giornate di studio in ricordo di Alessandra Concaro. Atti del seminario*, Torino, 2011, 39 ss.; A. Morelli, *Doveri costituzionali e principio di solidarietà*, in *Diritto costituzionale*, 2, 2019, 5 ss.

⁶⁷ Cfr. ad esempio A. Spadaro, *Dall'indisponibilità (tirannia) alla ragionevolezza (bilanciamento) dei diritti fondamentali. Lo sbocco obbligato: l'individuazione di «doveri» altrettanto fondamentali*, in *Politica del diritto*, 1, 2006, 167; L. Nazzicone, *I doveri "presi sul serio"*, in *Foro Italiano*, 5, 2020, 293 ss.

Repubblica⁶⁸. Allo stesso tempo, la migliore dottrina ha da tempo fatto notare che i doveri di solidarietà previsti dall'art. 2 rappresentano un presupposto indissolubile della concreta realizzazione dei diritti fondamentali, permettendo a quest'ultimi di acquisire concretezza anche nei rapporti tra privati⁶⁹. In questo senso, dunque, pare pienamente da sposare l'orientamento che propende per una rinnovata valorizzazione della categoria.

Il tema dei doveri di solidarietà acquista specifica rilevanza di fronte all'avvento dell'intelligenza artificiale, e, più in generale, di fronte alla c.d. società del rischio, frutto della professionalizzazione, dovuta al progresso tecnico, di pressoché ogni settore produttivo⁷⁰. Lo sviluppo economico e tecnologico degli ultimi decenni, infatti, è andato di pari passo al diffondersi di attività ad alta complessità tecnica e intrinsecamente pericolose, ma comunque consentite in ragione del benessere sociale che ne deriva. Ciò ha portato all'elaborazione di articolate procedure di documentazione e gestione del rischio, e alla creazione di un numero crescente di obblighi di prevenzione in capo agli esperti, spesso presidiati anche dal punto di vista penale⁷¹. Anche il settore dell'intelligenza artificiale dovrà, prevedibilmente, affrontare una simile evoluzione, e la comunità degli esperti, in adempimento dei doveri di solidarietà derivanti dal possesso di conoscenze specialistiche, sarà gravata della responsabilità di orientarne lo sviluppo nel senso dell'affidabilità e della centralità dell'essere umano, anche attraverso penetranti forme di collaborazione col decisore politico, che dovrà dimostrarsi all'altezza del compito. Allo stesso tempo, sulle istituzioni statali graverà l'obbligo già menzionato di predisporre adeguati programmi di istruzione, in modo da incrementare le conoscenze di base in materia di intelligenza artificiale e favorire la formazione di un numero crescente, e sempre più necessario, di esperti.

9. Alcune conclusioni

Dall'analisi fin qui condotta, emerge come l'intelligenza artificiale ponga questioni urgenti per i diritti fondamentali da due distinti punti di vista: in

⁶⁸ Sull'inidoneità del dovere di fedeltà alla Repubblica a bilanciare i diritti di libertà v. R. Bin, G. Pitruzzella, *Diritto costituzionale (XXII ed.)*, 2021, 600 ss.

⁶⁹ Si vedano *ex multis* G. M. Lombardi, *Contributo allo studio dei doveri*, cit.; S. Rodotà, *La vita e le regole. Tra diritto e non diritto*, Milano, 2006, 39 ss.; F. Polacchini, *Doveri costituzionali e principio di solidarietà*, Bologna, 2016.

⁷⁰ Il riferimento principale è a U. Beck, *La società del rischio. Verso una seconda modernità*, Francoforte, 1986.

⁷¹ Con la conseguente espansione della responsabilità colposa, estensivamente commentata dalla dottrina penalistica, cfr. ad es. i contributi raccolti in G. De Francesco, G. Morgante (a cura di), *Il diritto penale di fronte alle sfide della "società del rischio"*, Torino, 2017; M. Catino, *Errori e disastri nei sistemi complessi*, in *Diritto Penale e Uomo*, 2, 2020, 1 ss; e le riflessioni raccolte, a partire dal 2019, in Aa. Vv., *Il diritto penale nella società del rischio*, "cantiere aperto" della Rivista *Diritto Penale e Uomo*, accessibile al link: bit.ly/3BwNJdJ.

primo luogo, ponendo sfide inedite per la tutela dei diritti già riconosciuti nella nostra tradizione costituzionale; in secondo luogo, rendendo necessario elaborare *nuovi diritti*, con cui proteggere l'individuo dai rischi che tale rivoluzione tecnologica porta con sé.

Il diritto ad essere resi consapevoli dell'utilizzo di tecnologie intelligenti, la prima delle nuove posizioni giuridiche prese in esame, pare il presupposto irrinunciabile per uno sviluppo tecnologico armonico e in cui le esigenze dell'essere umano rimangano al centro. Il diritto alla spiegazione e il diritto al controllo umano, gli altri due nuovi diritti trattati, sono strettamente legati al requisito tecnico della trasparenza del sistema, la cui realizzabilità, come estensivamente detto, varia di molto a seconda delle applicazioni dell'intelligenza artificiale coinvolte. Saranno gli interessi di volta in volta in gioco, così, a determinare quale sia il grado di trasparenza minimo desiderabile dei sistemi coinvolti, e, dunque, l'estensione dei due diritti in esame all'esito del bilanciamento. Si auspica, allora, che la loro eventuale positivizzazione avvenga con norme flessibili e aperte allo sviluppo tecnologico, adatte a tutelare l'individuo senza imporre al mercato, ove non sia necessario, requisiti tecnici di difficile realizzazione. Come detto, la Proposta di Regolamento della Commissione Europea pare in larga misura idonea allo scopo, pur con alcune cautele in materia di *human oversight* (art. 14) che, in caso di approvazione dell'attuale formulazione, starà all'interprete risolvere.

Infine, l'avvento dell'intelligenza artificiale non è privo di conseguenze anche dal punto di vista dei doveri costituzionali. Infatti, spetterà innanzitutto agli esperti, in sinergia col decisore politico, indirizzarne lo sviluppo nella giusta direzione, in adempimento dei doveri di solidarietà derivanti dall'essere i depositari delle necessarie conoscenze tecniche.