

2/2026 anteprima

Rivista di diritto dei media

ISSN 2532-9146

Decisioni automatizzate e autonome: il diritto alla spiegazione e la sfida della sua effettività nello Stato di diritto

Gianluca Fasano

Sommario

1. Introduzione. – 2. Il diritto alla spiegazione come strumento di tutela della persona e garanzia effettiva dello Stato di diritto. – 3. Automazione e autonomia: premesse per il diritto alla spiegazione. – 4. Il diritto alla spiegazione nel GDPR: dal dibattito dottrinale alla giurisprudenza europea. – 5. L’approccio sistematico dell’AI Act in termini di interpretabilità e spiegazione. – 6. Contenuto e limiti del diritto alla spiegazione nell’AI Act. – 7. Conclusioni.

1. Introduzione

Utilizziamo sempre più spesso artefatti non soltanto per facilitare il compimento di un’azione, un’attività o un’operazione già pianificata, ma anche per affidare loro compiti di valutazione, previsione e decisione secondo modalità automatizzate o autonome. Man mano che la tecnologia continuerà ad evolversi, questi artefatti sono destinati a integrarsi sempre più profondamente nelle nostre vite e nelle pratiche decisionali, producendo non soltanto una trasformazione profonda dei rapporti di potere all’interno della società ma anche riflessi rilevanti sulla relazione tra tecnica e persona¹.

¹ Osserva S. Mangiameli, *Verità e certezza nell’esperienza costituzionale*, in *Diritto e Società*, 4, 2024, 643: «Il punto in cui si genera il diritto è la “persona”, con quel che l’espressione significa, e si genera come portato di questa e non come frutto di un utile, di una forma o dello Stato». Circa il ruolo della politica nel costruire il rapporto tra tecnica e diritto, v. *ex multis* G. De Minico, *Nuova tecnica per nuove disegnanze. Case law: Disciplina Telecomunicazioni, Digital Services Act e Neurodiritti*, in *federalismi.it*, 6, 2024, 2, ove l’Autrice enfatizza il ruolo della politica nel rapporto tra tecnica e diritto, precisando che: «La tecnica intrattiene un dialogo continuo col diritto, il cui ritmo, ora accelerato ora rallentato, dipende dalla

L'avvento dei sistemi decisionali automatizzati² è dovuto allo sviluppo di macchine dotate di ampie capacità di calcolo applicate a enormi quantità di dati e ha rivoluzionato numerosi aspetti della nostra società, influenzando in maniera significativa le modalità con cui vengono prese decisioni cruciali nella e per la vita di tutti i giorni. Tali sistemi vengono utilizzati in una vasta gamma di settori, dove è necessario analizzare una grande quantità di informazioni, non soltanto di carattere personale, al fine di valutare, classificare, prevedere e prendere decisioni³.

Si pensi al settore bancario e finanziario, ove sono diffusi per valutare il rischio di credito di un individuo o di un'azienda, indirizzando la decisione circa la concessione di un prestito, il tasso di interesse e le relative condizioni⁴. Impieghi se ne trovano anche nel settore assicurativo, ove le compagnie possono utilizzarli per determinare i premi assicurativi, nel settore immobiliare, dove i proprietari possono valutare la solvibilità dei potenziali inquilini, nel mercato del lavoro, in cui i datori di lavoro possono utilizzarli nel processo di screening dei candidati, sia per il reclutamento sia per la progressione professionale⁵, e perfino nel settore pubblico, ove possono essere impiegati per decidere l'accesso a prestazioni o servizi pubblici e le relative modalità di erogazione⁶. Si tratta soltanto di alcuni dei numerosi

volontà del legislatore di scegliere se e in che misura piegare il fattore tecnico a un obiettivo a sé gradito».

² Nel seguito di questo lavoro, *infra* par. 3, verrà precisato il rapporto di genere a specie tra decisioni automatizzate e autonome.

³ Per le tematiche inerenti al rapporto tra nuove tecnologie e protezione dei dati personali, si veda G. Cerrina Feroni (a cura di), *Il ruolo del Garante per la protezione dei dati personali*, Bologna, 2023.

⁴ La Fair Isaac Corporation (FICO) è una società statunitense specializzata nell'analisi del rischio creditizio. Nel 1989 introdusse il FICO Score, destinato a divenire uno degli strumenti standardizzati più diffusi per la valutazione del merito creditizio negli Stati Uniti. Il ricorso a informazioni creditizie per l'assunzione di decisioni economiche è, tuttavia, assai precedente all'era dei Big Data, come testimoniato dal *Fair Credit Reporting Act*, emanato nel 1970 negli Stati Uniti per promuovere «l'accuratezza, l'equità e la privacy delle informazioni» sui rapporti di credito, secondo la Federal Trade Commission.

⁵ Per un'analisi sulle discriminazioni prodotte da software impiegati nelle procedure di assunzione, v. A. Kelly-Lyth, *Challenging biased hiring algorithms*, in *Oxford Journal of Legal Studies*, 4, 2021, 899 ss.

⁶ Sull'uso di sistemi decisionali automatizzati nel settore dei pubblici servizi è noto il caso di SyRI (System Risk Indication), un sistema progettato nei Paesi Bassi in grado di analizzare grandi quantità di dati e di ottenere come risultato un indicatore di rischio, cioè una stima della probabilità che una persona fisica o giuridica possa usare illecitamente, ovvero in modo non conforme, i sussidi pubblici. Il Tribunale distrettuale dell'Aia ha ritenuto la disciplina normativa di SyRI incompatibile con l'art. 8 CEDU, in quanto l'ingerenza nella vita privata non rispettava il giusto equilibrio richiesto tra gli obiettivi perseguiti dallo Stato e la tutela della vita privata degli interessati (Rechtbank Den Haag, 5 febbraio 2020, C-09-550982-HA 18-388). L'interferenza esercitata dallo Stato nella vita

ambiti nei quali l'automazione è ormai entrata nei processi decisionali. Dunque, se il diritto «si genera come portato [...della persona...] e non come frutto di un utile, di una forma o dello Stato»⁷, allora la progressiva attribuzione a sistemi automatizzati di funzioni valutative e decisionali tradizionalmente affidate alla mediazione umana non rappresenta solo un cambiamento tecnico, ma investe in modo nuovo il modello costituzionale dello Stato di diritto, in cui la persona è origine e fine dell'ordine giuridico⁸ e in cui le libertà fondamentali vanno protette in modo effettivo rispetto alle forme di potere, pubblico o privato⁹.

Per tali ragioni, il diritto europeo ha mostrato un crescente e progressivo interesse verso le decisioni automatizzate, dapprima¹⁰ con una disciplina specifica per le decisioni computazionali alimentate da dati personali (regolamento (UE) 2016/679, concernente la protezione dei dati personali, “GDPR”)¹¹, successivamente con un quadro normativo orizzontale che, riferito ai sistemi di intelligenza artificiale e ai rischi derivanti dal loro impiego, può applicarsi anche ai sistemi decisionali che trattino dati personali (regolamento (UE) 2024/1689 sull'intelligenza artificiale, “AI Act”)¹².

privata non è stata giudicata proporzionata rispetto allo scopo perseguito in una società democratica.

⁷ Così S. Mangiameli, *Verità e certezza nell'esperienza costituzionale*, cit., 643.

⁸ Il principio personalista affermato dalla Costituzione italiana non si limita a riconoscere la persona come origine e fine dell'ordinamento ma afferma la persona nella sua realtà sociale e relazionale, tant'è che «la Repubblica riconosce e garantisce i diritti inviolabili dell'uomo sia come singolo sia nelle formazioni sociali ove si svolge la sua personalità, [...]» (art. 2 Cost.). E per comprendere senso e portata di questa disposizione occorre considerare «il clima storico-culturale dell'Assemblea Costituente, frutto dell'incontro (e dell'accordo) di tre diverse ispirazioni di pensiero: quello liberale, quello cattolico e quello socialista. La discussione in Assemblea Costituente sull'art. 2, infatti, fu particolarmente sentita e vide emergere istanze giusnaturalistiche, espressione del personalismo di matrice cattolica, istanze giuspositiviste proprie della cultura di sinistra, che sottolineavano il ruolo dello Stato nel rendere efficaci i diritti, e istanze garantistiche di matrice liberale, orientate a sottolineare la dimensione individualista dei diritti», così S. Mangiameli, *I diritti costituzionali: dallo Stato ai processi di integrazione*, Torino, 2020, 176.

⁹ Per un approfondimento sui rapporti tra tecnica, diritto e democrazia in Italia, v. G. Cantarano (a cura di), *Tecnica, politica, democrazia*, Napoli, 1994; E. Severino, *Democrazia, tecnica, capitalismo*, Brescia, 2009; G. De Minico, *Algoritmi tra regolazione e democrazia*, in *Rivista critica del diritto privato*, 2019, 345 ss.

¹⁰ L'art. 15 della direttiva 95/46/CE già riconosceva il diritto a non essere sottoposti a determinate decisioni fondate esclusivamente su trattamento automatizzato.

¹¹ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati).

¹² Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i

L'evoluzione più evidente del diritto europeo sta nel fatto che il GDPR costruisce la spiegazione come parte di una posizione giuridica complessa, comprensiva di partecipazione e contestazione, mentre l'AI Act configura soprattutto un diritto informativo strumentale al controllo delle decisioni ma non autosufficiente rispetto alla sua contestazione. Queste differenze strutturali discendono da un preciso approccio normativo che, tenendo conto dello sviluppo delle macchine autonome, interviene in modo sistematico a consolidare il diritto alla spiegazione come strumento informativo ma complementare agli ulteriori rimedi previsti dall'ordinamento.

Nelle prossime pagine si analizzerà tale evoluzione normativa al fine di verificare se, e in quali termini, l'approccio sistematico seguito dal diritto europeo possa restituirci un livello di effettività della spiegazione tale da garantire la tutela della persona a fronte dell'esercizio del potere decisionale algoritmico. In effetti, l'introduzione di siffatte garanzie risponde all'esigenza di adattamento dell'ordinamento alle trasformazioni tecnologiche. Eppure, essa deve fare i conti con un paradosso: proprio i recenti sviluppi tecnologici, in particolare le caratteristiche che rendono i sistemi automatizzati efficienti — velocità, volume di dati elaborati, capacità di individuare correlazioni non immediatamente percepibili dall'uomo — possono rendere più difficile la comprensione del processo che conduce a un determinato risultato, soprattutto nei sistemi più complessi e che operano su larga scala. È il paradosso della razionalità automatizzata che, proprio nel contesto decisionale, dispiega tutta la sua forza problematica. Occorrerà, dunque, verificare se il diritto alla spiegazione delle decisioni automatizzate e autonome sia sufficientemente strutturato e determinato, al fine di consentire alla persona di poter attivare in modo effettivo le garanzie proprie dello Stato di diritto.

regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale – AI Act). L'interesse del legislatore europeo per un quadro regolatorio volto a minimizzare i rischi collegati all'impiego dell'AI si è sviluppato lungo un percorso che ha conosciuto anche l'emaneazione di numerosi atti di *soft law*, tra cui si ricordano gli *Orientamenti etici per un'IA affidabile*, del Gruppo di esperti ad alto livello sull'intelligenza artificiale istituito dalla Commissione europea nel giugno 2018. Tra le posizioni espresse in dottrina sul tema si veda, *ex multis*, 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, 422; E. Stradella, *La regolazione della Robotica e dell'Intelligenza artificiale: il dibattito, le proposte, le prospettive. Alcuni spunti di riflessione*, in questa *Rivista*, 1, 2019, 77-78; G. Resta, *Cosa c'è di 'europeo' nella Proposta di Regolamento UE sull'intelligenza artificiale?*, in *Diritto dell'informazione e dell'informatica*, 2, 2022, 323 ss.

2. Il diritto alla spiegazione come strumento di tutela della persona e garanzia effettiva dello Stato di diritto

La relazione tra l'impiego di sistemi decisionali e la concezione dello Stato di diritto emerge con particolare evidenza considerando la diffusione dei sistemi decisionali automatizzati nella gestione di aspetti cruciali della vita delle persone: si pensi all'accesso al credito, all'accesso e gestione del lavoro, all'erogazione di servizi pubblici e privati, alla distribuzione di risorse, ecc. L'elemento di maggiore rilevanza di tale diffusione non risiede solo nella moltitudine degli ambiti interessati bensì nella qualità del mutamento che esso determina: nei sistemi ad architettura complessa, l'intermediazione computazionale tende a orientare la discrezionalità umana mediante modelli predittivi e criteri statistico-probabilistici. Il rischio è che tali sistemi spesso non siano intelligibili dai destinatari delle decisioni, finendo per comprimere il diritto alla difesa, al contraddittorio e, nei provvedimenti pubblici, il diritto alla motivazione. In definitiva, il problema assume una diretta rilevanza per lo Stato di diritto, poiché ciò che viene messo in discussione è la controllabilità stessa della decisione.

Questo scenario rappresenta un vero paradosso della razionalità automatizzata. Per un verso, le nuove tecnologie hanno il vantaggio di espandere i confini della cognizione umana, determinando la "liberazione" della persona dai propri limiti cognitivi e assicurando alla società intera una profonda innovazione¹³. Applicazioni virtuose le ritroviamo negli ambiti in cui i benefici sono rivolti alla collettività, come nella ricerca scientifica e nel campo biomedico, in quello della sicurezza e le dovremmo trovare, sempre più diffusamente, nell'ambito della pubblica amministrazione¹⁴.

¹³ La visione protetica venne teorizzata oltre un ventennio fa da V. Frosini, *Il giurista e le tecnologie dell'informazione*, Roma, 1998, 58, il quale evidenziava come l'informatica avesse determinato una vera mutazione antropologica: «Dialogare attraverso le macchine, o con le macchine stesse, non è la stessa cosa di conversare a viva voce: si adopera una protesi intellettuale, che modifica il procedimento del conoscere e del comunicare. L'uomo è divenuto essenzialmente diverso da quello che è stato in tutti i tempi, e non potrebbe tornare indietro senza perdere la sua nuova umanità, che gli è stata conferita dalla cultura informatica di massa».

¹⁴ In tale prospettiva si colloca anche quanto previsto dall'art. 7 della legge italiana sull'intelligenza artificiale (n. 132/2025), che in materia di uso dell'intelligenza artificiale in ambito sanitario e della disabilità stabilisce che «L'utilizzo di sistemi di intelligenza artificiale contribuisce al miglioramento del sistema sanitario, alla prevenzione, alla diagnosi e alla cura delle malattie, nel rispetto dei diritti, delle libertà e degli interessi della persona, anche in materia di protezione dei dati personali». Benché la norma sia comprensibilmente orientata ad assicurare l'efficienza dei servizi unitamente alla centralità della persona, assume un tono molto programmatico e poco pragmatico, configurandosi come una disposizione di indirizzo, volta a fissare un quadro di riferimento valoriale e

Nel campo delle decisioni automatizzate, in particolare, le ragioni per sostituire gli esseri umani sono solitamente ricondotte ai vantaggi derivanti dall'elaborazione dei dati su larga scala, dalla velocità, dall'elevato volume di informazioni trattabili e alle aspettative di tassi di errore inferiori rispetto a quelli degli esseri umani¹⁵. Alle macchine non sfugge nulla, loro non si stancano¹⁶. Viene riconosciuta loro una maggiore efficienza rispetto alle capacità analitiche umane: gli algoritmi predittivi, ad esempio, sono in grado di analizzare quantità elevatissime di dati, anche fra loro eterogenei sotto il profilo qualitativo, individuando preziose correlazioni, rapporti, inferenze che nessuna valutazione umana sarebbe in grado di rilevare¹⁷.

Per altro verso, queste stesse caratteristiche di funzionamento dei sistemi decisionali automatizzati più complessi, nei quali l'automazione può accompagnarsi a capacità inferenziali e a differenti gradi di autonomia rispetto all'intervento umano, possono sfuggire alla comprensione analitica dell'uomo. Seppur tale capacità rappresenti un notevole vantaggio in quei contesti in cui sono essenziali la rapidità decisionale, l'adattamento alla complessità e la risposta all'imprevedibilità, nell'ambito delle decisioni automatizzate che incidono sulla sfera giuridica della persona essa manifesta il pericolo di decisioni dagli esiti imprevedibili, con conseguenze negative in termini di verifica delle decisioni assunte¹⁸.

Tale aspetto diviene ancora più marcatamente problematico per due ordini

sistematico per l'interpretazione e l'attuazione di norme prescrittive.

¹⁵ In termini di riduzione dell'errore umano si è espressa di recente anche l'Assemblea Parlamentare del Consiglio d'Europa, Risoluzione 2628 del 3 ottobre 2025, su "*Artificial intelligence and migration*", al cui punto 5 si legge: «*in some cases AI may offer greater security and effectiveness than human decision making alone, by reducing the scope for human error*» (enfasi aggiunta).

¹⁶ Cfr. OECD, *Artificial Intelligence in Society*, Paris, 2019, ove si evidenzia che le tecnologie di intelligenza artificiale promettono di generare guadagni di produttività, migliorare il benessere e aiutare ad affrontare le sfide globali, come il cambiamento climatico, la scarsità di risorse e le crisi sanitarie.

¹⁷ Cfr. Gruppo di lavoro Articolo 29 per la protezione dei dati, *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679*, adottate il 3 ottobre 2017 (versione emendata e adottata il 6 febbraio 2018) (WP 251 rev.01), 5, ove si legge che «La profilazione e il processo decisionale automatizzato possono essere utili per le persone fisiche e le organizzazioni, offrendo loro vantaggi quali: miglioramenti dell'efficienza; risparmi di risorse. Presentano inoltre numerose applicazioni commerciali: ad esempio, possono essere utilizzati per segmentare meglio i mercati e personalizzare i servizi e i prodotti allineandoli alle singole esigenze. Anche la medicina, l'istruzione, l'assistenza sanitaria e i trasporti possono beneficiare di questi processi».

¹⁸ Per una panoramica sull'intelligibilità delle decisioni macchiniche si veda F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge (Mass.), 2015.

di ragioni: da un lato, per via del passaggio dalla fase dell'automazione a quella dell'autonomia¹⁹, di cui si parlerà meglio in seguito; dall'altro, soprattutto poiché il potere (algoritmico) è spesso azionato da soggetti privati²⁰. In tale scenario, si registra una tensione evidente tra accrescimento della razionalità tecnica e contrazione di quella giuridica, tanto da chiamare in causa l'architettura dello Stato di diritto. Il suo ruolo non è certamente quello di frenare l'innovazione tecnologica, ma di incanalarla secondo principi costituzionali, affinché l'efficienza cui mira l'automazione decisionale non prevalga sulla tutela della persona. Il diritto alla spiegazione tende a questo fine²¹.

L'evoluzione della concezione dello Stato di diritto mostra che la governance dell'algoritmo si iscrive in piena continuità con essa: ogniqualvolta nuove forme di potere, pubblico o privato, acquisiscano la capacità di incidere in modo significativo sulla sfera giuridica della persona, emergono garanzie volte a razionalizzare l'esercizio del potere²². Un'ulteriore evoluzione delle forme di garanzia proprie dello Stato di diritto si è avuta grazie al contributo della dottrina del c.d. *Drittwirkung*, in forza della quale oggi non si dubita che le garanzie costituzionali, originariamente a presidio dei rapporti verticali tra individuo e pubblici poteri²³, trovino applicazione an-

¹⁹ Per A. Simoncini, *Il linguaggio dell'intelligenza artificiale e la tutela costituzionale dei diritti*, in *Rivista AIC*, 2, 2023, 1 ss., spec. 12, «Alla macchina non si chiede più di eseguire ciò che il soggetto ha deciso, ma le si chiede di decidere: è la distinzione tra “automazione” ed “autonomia”». Al tema della discontinuità tecnologica tra “automazione” e “autonomia” e alle sue implicazioni per la *rule of law* riserva particolare attenzione O. Pollicino, *Regolazione e innovazione tecnologica nell'ordinamento della rete*, in *Rivista AIC*, 2, 2025, 119 ss.

²⁰ Sul punto, è molto chiara la posizione di G. De Minico, *Nuova tecnica per nuove disuguaglianze*, cit., 3, ove scrive che «Le rivendicazioni giacobine verso il potere costituito sono diventate le rivendicazioni digitali dei *netizens* nei confronti delle Autorità private, e, in seconda battuta, anche verso il potere pubblico, perché nell'ecosistema digitale i privati *well founded and well structured* sono i nuovi tiranni delle libertà, potendole stracciare o appropriarsene, come un tempo faceva lo Stato».

²¹ Il differente concetto di spiegabilità dell'AI è stato molto enfatizzato anche sul piano etico, si pensi agli *Orientamenti etici per un'IA affidabile*, del Gruppo di esperti ad alto livello sull'intelligenza artificiale istituito dalla Commissione europea nel giugno 2018, *Artificial Intelligence, Ethics Guidelines for Trustworthy AI* (2019).

²² Per C. Schmitt, *Dottrina della Costituzione* (1928), trad. it., Milano, 1984, 171 ss., il concetto di Stato di diritto si lega al riconoscimento dei diritti fondamentali e alla tutela giurisdizionale di questi nei confronti dello Stato.

²³ In G. Jellinek, *La dichiarazione dei diritti dell'uomo e del cittadino (Die Erklärung der Menschen- und Bürgerrechte)*, III ed., 1909, I ed. 1895), trad. it. D. Nocilla, Milano, 2002, 104, la nascita dei diritti di libertà viene così spiegata: «La pressione che i poteri dominanti esercitarono sulla libertà di determinarsi dell'individuo creò l'idea che uno speciale diritto corrispondesse a ciascuna delle direzioni in cui si esercitava l'oppressione. Così nasce, oltre alla rivendicazione della libertà religiosa, quella della libertà di stampa, della libertà di parola, della libertà di associazione e riunione, della libertà di espatrio, del diritto di

che nei rapporti orizzontali, tra soggetti privati²⁴. Ne deriva, così, un assetto ordinamentale nel quale la tutela delle libertà fondamentali assume una portata generalizzata.

Ora, se lo Stato di diritto nasce per delimitare l'esercizio arbitrario del potere attraverso la predisposizione di norme giuridiche e si fonda sulla centralità della persona e sulla garanzia dei suoi diritti fondamentali²⁵, allora un sistema computazionale che sia posto in condizione di produrre decisioni che incidono su diritti e libertà assume inevitabilmente una rilevanza costituzionale²⁶. L'irruzione dell'automazione decisionale, anche nella forma dell'AI, rappresenta un ulteriore banco di prova per lo Stato di diritto. Il c.d. 'costituzionalismo digitale'²⁷, inteso come «una costellazione

petizione, della libertà dall'arresto».

²⁴ In fase di elaborazione e riconoscimento dei diritti fondamentali l'attenzione venne rivolta all'esigenza di proteggere la persona dalle ingerenze dei pubblici poteri. Successivamente, l'attenzione è via via cominciata a spostarsi sul diverso profilo della tutela della persona nell'ambito dei rapporti tra individui, cioè nei rapporti tra privati. Per ripercorrere la dottrina tedesca sull'applicabilità diretta delle norme costituzionali nei rapporti orizzontali (privato/privato), elaborando la figura della c.d. *Drittwirkung*, si veda P. Femia (a cura di), *Drittwirkung: principi costituzionali e rapporti tra privati*, Napoli, 2018. C'è chi ritiene che la *Drittwirkung* rappresenti solo un falso problema, poiché a tutti i soggetti, pubblici e privati, si impongono le norme costituzionali; v. R. Manfellotti, *Diritto pubblico*, Torino, 2020, 70. In giurisprudenza, si veda Corte costituzionale, 26 luglio 1979, n. 88, considerato in diritto n. 2, in cui la Corte riconosce che il bene alla salute «è tutelato dall'art. 32 Costituzione non solo come interesse della collettività, ma anche e soprattutto come diritto fondamentale dell'individuo, sicché si configura come un diritto primario ed assoluto, pienamente operante anche nei rapporti tra privati». *Conforme* Corte costituzionale, 14 luglio 1986, n. 184, considerato in diritto n. 11.

²⁵ Cfr. S. Rodotà, *Il diritto di avere diritti*, Roma-Bari, 2012, 401-402, ove l'Autore evidenzia come il fenomeno riguardi scelte e decisioni sia a livello individuale, nella vita quotidiana delle persone, sia a livello collettivo, all'interno della società nel suo complesso.

²⁶ Sul tema si veda S. Mangiameli, *Verità e certezza nell'esperienza costituzionale*, in *Diritto e Società*, 4, 2024, 653, ove l'Autore enfatizza che «La mirabile costruzione dello "Stato di diritto", fatta di equilibri e composta di elementi ben noti alla scienza del diritto costituzionale, quali i diritti e le libertà, la divisione dei poteri, la democrazia e l'autonomia delle formazioni sociali (corporative e territoriali) può sempre essere messa in discussione e il tempo presente ne è proprio l'esempio». Si veda, tra gli altri, V.P. Costa-D. Zolo, *Lo Stato di diritto. Storia, teoria, critica*, Milano, 2006.

²⁷ In tali termini A. Lamberti, *Costituzionalismo digitale, poteri delle piattaforme*, in *Consulta online*, 2, 2025, 1084, ove si precisa che «il costituzionalismo digitale nasce a fronte dell'esigenza di trovare adeguata soluzione al "problema" costituzionale dell'emersione dei poteri privati, con specifico riferimento alle grandi piattaforme digitali». Sul tema del potere privato proprio delle piattaforme online, cfr., *ex multis*, M. Bassini, *Internet e libertà di espressione. Prospettive costituzionali e sovranazionali*, Roma, 2019, 107 ss. e O. Grandinetti, *Facebook vs. CasaPound e Forza Nuova, ovvero la disattivazione di pagine social e le insidie della tutela multilivello dei diritti fondamentali*, in questa *Rivista*, 2021, 173 ss. V. anche A. Papa, *Espressione e diffusione del pensiero in Internet. Tutela dei diritti e progresso tecnologico*, Torino, 2009, spec. 241 ss. Tra le più recenti trattazioni monografiche, cfr. F. Paruzzo, *I sovrani della rete. Piattaforme*

di iniziative che hanno provato ad articolare un insieme di diritti politici, norme di governance, e limitazioni all'esercizio del potere su Internet»²⁸, è chiamato ad affrontare le sfide poste da una tecnologia che non si limita a supportare l'esecuzione di decisioni umane, ma assume un ruolo attivo nella produzione di decisioni, in alcuni casi attraverso sistemi qualificati come autonomi, che incidono direttamente sulle persone. È la stessa struttura della decisione a esser trasformata, divenendo il portato di una relazione uomo-macchina che si articola e sviluppa di continuo, sin dalla scelta dei dati da assumere come input del sistema e fino al momento dell'assunzione stessa della decisione. Peraltro, le decisioni algoritmiche si fondano su processi probabilistici e statistici piuttosto che su ragionamenti logico-deduttivi tradizionali, facendo emergere questioni complesse rispetto alla possibilità, per la persona, di sottrarsi alle proiezioni elaborate sulla base dei propri comportamenti pregressi²⁹.

In questa prospettiva, in cui la possibile sottrazione della decisione alla controllabilità pone un problema allo Stato di diritto, il diritto alla spiegazione non si esaurisce nella pretesa di comprendere analiticamente il funzionamento interno delle logiche computazionali, ma va inteso come strumento giuridico volto a preservare la controllabilità dell'agire mediante computazione. La questione che si pone al costituzionalismo contemporaneo è, allora, se e a quali condizioni il diritto alla spiegazione sia effettivamente in grado di ricondurre l'esercizio del potere mediato da tali sistemi entro le esigenze di controllabilità proprie dello Stato di diritto, a garanzia della centralità della persona³⁰.

digitali e limiti costituzionali al potere privato, Napoli, 2022; E. Cremona, *I poteri privati nell'era digitale. Libertà costituzionali, regolazione del mercato, tutela dei diritti*, Napoli, 2023.

²⁸ L. Gill-D. Redeker-U. Gasser, *Towards Digital Constitutionalism? Mapping Attempts to Craft an Internet Bill of Rights*, Berkman Center Research Publication No. 2015-15, 2. In dottrina è stato proposto anche il concetto di «cyber-costituzionalismo» che prevenga gli abusi e garantisca «che l'espansione del digitale avvenga nel rispetto della dignità umana, dell'uguaglianza e della giustizia sociale, contribuendo a costruire una società virtuale sostenibile e ben regolata», così M.C. Canato, *La tutela dei diritti nel metaverso*, in F. Casa-S. Gaetano-G. Pascali (a cura di), *Intelligenza artificiale: diritto, etica, democrazia*, Bologna, 2025, spec. 98-99.

²⁹ Nei sistemi decisionali fondati sulla profilazione, le persone vengono valutate sulla base di ciò che hanno fatto in passato e, secondo un calcolo di probabilità, da quel passato si inferisce ciò che presumibilmente faranno in futuro.

³⁰ Come scrive S. Mangiameli, *Verità e certezza nell'esperienza costituzionale*, cit., 644, «La Costituzione, curandosi della comunità politica, con il catalogo dei diritti e delle libertà e la distribuzione dei poteri pubblici tra gli organi dello Stato, per la quale nessuno di questi è sovrano e assoluto, testimonia questa costruzione dell'ordinamento che deriva dal principio personalista».

3. Automazione e autonomia: premesse per il diritto alla spiegazione

Per comprendere se il diritto alla spiegazione assicuri in modo effettivo le garanzie dello Stato di diritto, occorre tener conto che nel campo delle decisioni automatizzate assumono crescente rilievo i sistemi di intelligenza artificiale³¹, caratterizzati da gradi di autonomia variabili che accentuano le problematiche connesse alla controllabilità effettiva³². Diviene allora necessario analizzare l'evoluzione del diritto europeo e porre in raffronto l'art. 22 GDPR, dedicato al "Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione" e l'art. 86 AI Act, dedicato al "Diritto alla spiegazione dei singoli processi decisionali". Prima di procedere con tale raffronto, è necessaria una premessa per ricostruire il contesto concettuale entro il quale collocare il diritto alla spiegazione, contesto nel quale si registra il passaggio dalle decisioni unicamente automatizzate, caratterizzate dal paradigma dell'*assenza dell'uomo* (art. 22 GDPR)³³, alle decisioni mediate dall'intelligenza artificiale, nelle quali il paradigma è quello dell'autonomia che caratterizza la *relazione uomo-macchina* (art. 86 AI Act).

Quanto alle decisioni automatizzate, il punto di riferimento è rappresentato dal GDPR³⁴. Anche se il GDPR non detta un'esplicita definizione di decisione automatizzata, essa può comunque ricavarsi in forza del principio ermeneutico fondamentale del diritto dell'Unione: la ricostruzione

³¹ Cfr. B. Vaassen, *AI, Opacity, and Personal Autonomy*, in *Philosophy & Technology*, 35, 2022; C. Prunkl, *Human Autonomy at Risk? An Analysis of the Challenges from AI*, in *Minds and Machines*, 34, 2024.

³² La capacità di agire autonomamente e di prendere decisioni deliberate viene definita anche "agency", *ex multis*, E. Himma, *Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent?*, in *Ethics and Information Technology*, 1, 2009, 19 ss.

³³ L'art. 22 GDPR assume come elemento qualificante l'assenza di un intervento umano significativo.

³⁴ A ben vedere, già il Consiglio d'Europa con la Convenzione n. 108 del 28 gennaio 1981, ratificata con legge 21 febbraio 1989, n. 98 (Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati di carattere personale) fissa l'esigenza di garantire ad ogni persona «il rispetto dei suoi diritti e delle sue libertà fondamentali, e in particolare del suo diritto alla vita privata, in relazione all'elaborazione *automatica* dei dati a carattere personale che la riguardano» (art. 1 - corsivo aggiunto). Anche se il riferimento è all'emergente elaborazione informatica dei dati, nella relazione esplicativa della Convenzione è chiaro lo scenario cui si va incontro, caratterizzato da «un'ulteriore crescita dell'elaborazione automatica dei dati in ambito amministrativo, anche grazie alla riduzione dei costi di elaborazione dei dati, alla disponibilità di dispositivi di elaborazione dati "intelligenti" e all'istituzione di nuove infrastrutture di telecomunicazione per la trasmissione dei dati» (1).

di una disposizione può discendere dall'analisi del «contesto in cui essa si inserisce nonché degli obiettivi e della finalità che persegue l'atto di cui essa fa parte»³⁵. E così, la definizione di decisione automatizzata può esser ricostruita in chiave sistematica a partire dalle altre definizioni fissate nel regolamento. In particolare, rileva la definizione di “trattamento” (art. 4, n. 2, GDPR), che include qualsiasi operazione sui dati personali, sia manuale che automatizzata, ma che non specifica cosa sia un trattamento automatizzato. Si può ritenere che l'automazione qualifichi la tecnologia attraverso la quale sono compiute le operazioni di trattamento dei dati personali, per cui il trattamento automatizzato può essere inteso come trattamento nel quale le operazioni sui dati personali sono eseguite, in tutto o in parte, mediante sistemi informatici o altri strumenti tecnologici, anziché esclusivamente attraverso attività umana. In punto di fatto, l'automazione esclude che ogni singola operazione sia compiuta dall'uomo, esclude un intervento umano continuo. Anche le decisioni possono esser basate su un trattamento automatizzato e, qualora si fondino in modo esclusivo sull'automazione, l'assenza dell'uomo attiva le garanzie dell'art. 22 GDPR³⁶. La “profilazione” (art. 4, n. 4, GDPR) rappresenta una forma specifica di trattamento automatizzato finalizzata a valutare o prevedere aspetti personali di un individuo. Non ogni profilazione produce necessariamente una decisione automatizzata.

L'art. 22, par. 1, del GDPR mira a dettare tutele rafforzate con riferimento alle decisioni basate unicamente sul trattamento automatizzato e che hanno un'incidenza significativa sulla persona. La norma considera “decisioni” non solo quegli atti che producono effetti giuridici che attingono l'interessato ma anche atti che incidono significativamente su di esso in modo analogo. Quest'elemento appare dirimente per qualificare un atto come decisionale. Esso deve dimostrare incidenza sulla sfera personale degli interessati (non solo giuridica). Ciò è confermato dal considerando 71, per il quale una decisione può «includere una misura» che abbia la capacità di produrre gli effetti sopra visti. Sulla base di tali considerazioni, la Corte di giustizia ha attribuito alla nozione di decisione un significato ampio, tanto da includere «diversi atti che possono incidere sulla persona interessata in vari modi»³⁷.

³⁵ Cfr. CGUE, C-579/21, *Pankki S.* (2023).

³⁶ Cfr. linee guida del Gruppo di lavoro Articolo 29 per la protezione dei dati, *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679*, cit., 8, ove si legge che: «Il processo decisionale esclusivamente automatizzato consiste nella capacità di prendere decisioni impiegando mezzi tecnologici senza coinvolgimento umano».

³⁷ Cfr. CGUE, C-634/21, *SCHUEFA Holding AG* (2023), § 46, «la nozione di “decisione” ai sensi dell'articolo 22, paragrafo 1, del RGPD può quindi includere, come rilevato dall'avvocato generale al paragrafo 38 delle sue conclusioni, diversi atti che possono

Possiamo dunque assumere, ai fini della presente ricostruzione, che la decisione automatizzata costituisca l'esito di un'elaborazione automatizzata di dati personali effettuata mediante algoritmi, software, sistemi informativi che non richiedano un intervento umano continuo e che si manifesti in una diversità di atti, tutti accomunati dalla qualità di incidere sulla persona in modi differenti³⁸. Le tutele rafforzate previste dall'art. 22 GDPR si applicano soltanto alle decisioni assunte senza un intervento umano significativo, secondo il paradigma dell'assenza dell'uomo.

Ai fini della presente disamina, si può sostenere che una decisione basata su trattamento automatizzato, esclusivo o meno che sia, è comunque *machine-based*, nella misura in cui c'è una componente computazionale/macchinica senza la quale non vi sarebbe automazione. Su questa premessa può essere affrontato il diverso concetto di autonomia dell'AI.

I progressi tecnologici e le prestazioni dei moderni artefatti, in particolar modo quelli legati ai sistemi di *machine learning*³⁹, *deep learning* e all'utilizzo delle *deep neural networks*⁴⁰, hanno contribuito a diffondere la forte suggestione, presente in vari ambienti di studio dell'AI, che la macchina possa dirsi autonoma. A ben vedere, il concetto di autonomia viene adoperato in modo non univoco, a volte col significato di autonomia tecnica (capacità operativa)⁴¹, altre con quello di autonomia adattiva (capacità di adattamen-

incidere sulla persona interessata in vari modi, tale nozione è sufficientemente ampia da ricomprendere il risultato del calcolo della solvibilità di una persona sotto forma di tasso di probabilità relativo alla capacità di tale persona di onorare impegni di pagamento in futuro» (§ 46).

³⁸ Il concetto giuridico è abbastanza ampio e rispecchia il principio di neutralità tecnologica della norma, secondo cui la tutela giuridica deve potersi applicare indipendentemente dagli strumenti tecnologici utilizzati. Un'espressione di tale principio può esser considerata il considerando 15 GDPR, per il quale «La protezione delle persone fisiche dovrebbe applicarsi sia al trattamento automatizzato che al trattamento manuale dei dati personali» e non dovrebbe dipendere dalle tecnologie impiegate». In argomento, v. ex multis M. Hildebrandt-L. Tielemans, *Data protection by design and technology neutral law*, in *Computer Law & Security Review*, 5, 2013, 509 ss.; A. Ojanen, *Technology Neutrality as a Way to Future-Proof Regulation: The Case of the Artificial Intelligence Act*, in *European Journal of Risk Regulation*, 4, 2025, 1440 ss.

³⁹ D. Lehr-P. Ohm, *Playing with the data: What legal scholars should learn about machine learning*, in *University of California Davis Law Review*, 2, 2017, 671; T. Wang-B. Li-M. Chen-S. Yu, *Machine Learning Empowered Intelligent Data Center Networking: Evolution, Challenges and Opportunities*, Singapore, 2023, 10.

⁴⁰ Sul funzionamento di tali sistemi si rinvia a V. Nath-S.E. Levinson, *Autonomous Robotics and Deep Learning*, Heidelberg-New York-Dordrecht-London, 2014; G. Sartor, *L'informatica giuridica e le tecnologie dell'informazione, Corso d'informatica giuridica*, Torino, 2016, 278 ss.; F. Vincenzi, *Intelligenza artificiale, machine learning, deep learning*, in G. Ziccardi-P. Perri (a cura di), *Tecnologia e diritto. Informatica giuridica*, vol. II, Milano, 2019, 397 ss.

⁴¹ In questo senso può esser intesa la posizione di L. Floridi-G. Sanders, *On the Morality of Artificial Agents*, in *Minds and Machines*, 3, 2004, 377, i quali distinguono tra *moral agency*

to), altre ancora con quello di autonomia morale (capacità intenzionale)⁴². La caratteristica, rilevante ai fini della presente ricostruzione, non consiste nell'autoapprendimento durante l'esecuzione, né nella mera imprevedibilità dell'output, ma nella crescente capacità di alcuni sistemi di produrre risultati con un grado di indipendenza dall'intervento umano e mediante processi inferenziali non integralmente predeterminati caso per caso. La complessità di queste tecnologie rende arduo, persino per i progettisti e i programmatori, comprendere i meccanismi che hanno portato a uno specifico *output*⁴³. La distanza rispetto agli algoritmi tradizionali caratterizzati dalla «ripetibilità (determinismo): dati gli stessi dati in input l'algoritmo

e *artificial agency*, identificando in quest'ultima forme di autonomia funzionale pur prive di intenzionalità: «*Artificial agents can be described as entities capable of performing actions in an environment, but without the mental states typically associated with moral agency*». Nel ricostruire il pensiero morale classico P. Polieri, *Eppur fa muovere! L'IA tra neo-agency, centralità del dato e riconfigurazione del tempo. Un approccio multiprospettico*, in *Ratio Iuris*, 29.12.2025, scrive: «L'atto, in pratica, è moralmente valutabile perché riconducibile a un soggetto dotato di intenzione e consapevolezza, e, se ben si valuta, anche nei modelli più sofisticati di etica delle conseguenze o della virtù il riferimento ultimo resta l'essere umano come centro di imputazione morale. L'IA, al contrario, rompe questo schema non perché sia in sé *malvagia o autonoma* in senso forte, ma perché introduce una forma di *agency 'operativa'* che non coincide con l'*agency 'morale'* tradizionale» (2-3).

⁴² R. Cucchiara, *L'intelligenza non è artificiale. La rivoluzione tecnologica che sta già cambiando il mondo*, Milano, 2021, 122, secondo cui «I sistemi che progettiamo stanno piano piano imparando ad apprendere dal mondo esterno, sono in grado di farsi proprie rappresentazioni interne di capire di ragionare e quindi di generare azioni. Sbagliano. E sono capaci di ricominciare sapendo come migliorare dai propri errori. Come facciamo noi dalla mattina alla sera. Come dovremmo fare se fossimo intelligenti». Critico verso l'attribuzione di un'intenzionalità alle macchine è J.J. Bryson, *Robots Should Be Slaves*, in Y. Wilks (a cura di), *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*, Amsterdam, 2010, 63 ss.

⁴³ La nota metafora della c.d. *black box* descrive in modo emblematico l'opacità che caratterizza il funzionamento degli algoritmi decisionali: «possiamo osservare i suoi ingressi e le sue uscite, ma non sappiamo come l'uno si trasformi nell'altro», F. Pasquale, *The Black Box Society. The Secret Algorithms that Control Money and Information*, cit. La complessità dei calcoli e la natura non lineare dei sistemi computazionali moderni rendono inutile o eccessivamente gravoso l'impiego di tecniche di *reverse engineering*. In tal senso, si veda F. Donati, *Intelligenza artificiale e giustizia*, in *Rivista AIC*, 1, 2020, 428. Sulla c.d. *black box problem* v., *ex multis*, T. Wischmeyer, *Artificial Intelligence and Transparency: Opening the Black Box*, in T. Wischmeyer-T. Rademacher (a cura di), *Regulating Artificial Intelligence*, Cham, 2020, 80-81; V. Dignum, *Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way*, Cham, 2019, 59 ss.; D. Pedreschi et al., *Meaningful Explanations of Black Box AI Decision Systems*, in *Proceeding of the AAAI Conference on Artificial intelligence*, 1, 2019, 9780 ss.; M. Ebers, *Regulating AI and Robotics: Ethical and Legal Challenges*, in M. Ebers-S. Navas (a cura di), *Algorithms and Law*, Cambridge, 2020, 445 ss.

deve fornire gli stessi risultati in output»⁴⁴, può essere significativa⁴⁵.

A fronte del significativo mutamento del panorama tecnologico anche il quadro normativo ha conosciuto un'evoluzione significativa. L'Unione europea, con il recente regolamento sull'intelligenza artificiale, ha affrontato il tema dell'autonomia decisionale in una prospettiva più organica e sistematica, muovendo non solo dall'ambizione di presentarsi come «un leader mondiale nello sviluppo di un'AI sicura, affidabile ed etica»⁴⁶, ma dall'esigenza di coniugare lo sviluppo e l'impiego dei sistemi di intelligenza artificiale con «un elevato livello di protezione degli interessi pubblici, quali la salute e la sicurezza e la protezione dei diritti fondamentali, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, come riconosciuti e tutelati dal diritto dell'Unione»⁴⁷.

⁴⁴ Cfr. G. Sartor, *L'informatica giuridica e le tecnologie dell'informazione*, Torino, 2016, 85, il quale indica le seguenti proprietà dell'algoritmo «Finitezza: l'algoritmo deve portare alla soluzione in un numero finito di passi. Generalità: l'algoritmo non risolve solo un problema, ma una classe di problemi. Non ambiguità: le istruzioni sono specificate univocamente, cosicché la loro esecuzione avviene sempre nello stesso modo, indipendentemente dall'esecutore. Ripetibilità (determinismo): dati gli stessi dati in input l'algoritmo deve fornire gli stessi risultati in output».

⁴⁵ La distinzione appena abbozzata è stata acquisita anche dal Cons. Stato, sez. III, 25 novembre 2021, n. 7891, che si è espresso nei suddetti termini con riferimento a una fattispecie riguardante una gara d'appalto per la fornitura di pacemaker e defibrillatori per enti operanti in campo sanitario. In particolare, i giudici di Palazzo Spada hanno riaffermato la rilevanza della distinzione enfatizzando il ruolo di un sistema di *machine learning* che non si limita «ad applicare le regole software e i parametri preimpostati (come fa invece l'algoritmo "tradizionale") ma, al contrario, elabora costantemente nuovi criteri di inferenza tra dati e assume decisioni efficienti sulla base di tali elaborazioni» (§ 9.1 in diritto).

⁴⁶ Considerando 8 AI Act: «il presente regolamento contribuisce all'obiettivo di promuovere l'approccio antropocentrico europeo all'IA ed essere un leader mondiale nello sviluppo di un'IA sicura, affidabile ed etica».

⁴⁷ Cfr. considerando 8 AI Act: «Si rende pertanto necessario un quadro giuridico dell'Unione che istituisca regole armonizzate in materia di IA per promuovere lo sviluppo, l'uso e l'adozione dell'IA nel mercato interno, garantendo nel contempo un elevato livello di protezione degli interessi pubblici, quali la salute e la sicurezza e la protezione dei diritti fondamentali, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, come riconosciuti e tutelati dal diritto dell'Unione. Per conseguire tale obiettivo, è opportuno stabilire regole che disciplinino l'immissione sul mercato, la messa in servizio e l'uso di determinati sistemi di IA, garantendo in tal modo il buon funzionamento del mercato interno e consentendo a tali sistemi di beneficiare del principio della libera circolazione di beni e servizi. Tali norme dovrebbero essere chiare e solide nel tutelare i diritti fondamentali, sostenere nuove soluzioni innovative e consentire un ecosistema europeo di attori pubblici e privati che creino sistemi di IA in linea con i valori dell'Unione e sbloccino il potenziale della trasformazione digitale in tutte le regioni dell'Unione». Per O.J. Gstrein-N. Haleem-A. Zwitter, *General-purpose AI regulation and the European Union AI Act*, in *Internet Policy Review*, 3, 2024, l'AI Act ha un potenziale significativo per diventare un punto di riferimento globale per la governance e la regolamentazione in questo settore

La definizione di intelligenza artificiale adottata dal regolamento tratta in maniera esplicita la questione dell'autonomia dei moderni artefatti, evidenziando al contempo la relazione concettuale tra automazione e autonomia. L'art. 3 definisce un sistema di AI come: «un sistema *automatizzato* progettato per funzionare con livelli di *autonomia* variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, *deduce* dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali» (corsivo aggiunto)⁴⁸. L'ambito di applicazione del regolamento sull'AI ruota attorno a questa fondamentale definizione, non applicandosi indistintamente a tutti i sistemi ma solo a quelli che soddisfano la definizione normativa⁴⁹.

Il riferimento a un sistema automatizzato, ossia *machine-based system* nella versione inglese, riconduce anzitutto l'intelligenza artificiale alla categoria dei sistemi computazionali fondati sul funzionamento di macchine. Particolarmente significativo, in tale direzione, è il considerando 12, il quale afferma che: «il termine “automatizzato” si riferisce al fatto che il funzionamento dei sistemi di AI prevede l'uso di macchine». Si tratta di un'affermazione che colpisce per la sua estrema semplicità e chiarezza. Eppure, il termine “macchine” è impiegato in un'accezione ampia e funzionale per indicare qualsiasi insieme di elementi tecnologici (computer, software, server, programmi, sensori, processori, ecc.) capaci di eseguire processi di

di importanza globale strategica. Tuttavia, il suo successo dipende da un'applicazione efficace, da una proficua cooperazione intraeuropea e internazionale e dalla capacità dell'UE di adattarsi al panorama dell'intelligenza artificiale in rapida evoluzione.

⁴⁸ Tale definizione si ispira a quella elaborata dall'Organization for Economic Cooperation and Development (OECD, *Explanatory memorandum on the updated OECD definition of an AI system*, OECD Artificial Intelligence Papers, n. 8, Parigi, 2024) e, con ciò, riflette un preciso intento di politica normativa, ossia la necessità di fissare una nozione di “sistema di IA” che sia coerente con i lavori internazionali in materia. Tale allineamento definitorio non ha valore formale ma mira a favorire la convergenza normativa a livello internazionale e l'accettazione diffusa del quadro regolatorio, facilitando cooperazione e standard tecnici in un contesto che necessita di una governance globale. Per una posizione critica v. S. Wachter, *Limitations and loopholes in the EU AI Act and AI Liability Directives: what this means for the European Union, the United States, and beyond*, in *Yale Journal of Law and Technology*, 3, 2024, 671 ss., in cui l'Autrice rileva che i forti sforzi di lobbying delle grandi aziende tecnologiche e degli Stati membri siano stati purtroppo in grado di indebolire gran parte dell'AI Act.

⁴⁹ La complessità del fenomeno da regolare, unitamente alla necessità di adottare una definizione che sia flessibilmente adeguata a coprire gli sviluppi tecnologici (c.d. *futureproof*), giustifica la previsione di orientamenti sull'applicazione del regolamento che, seppur non vincolanti, mirano a conciliare l'esigenza della certezza del diritto con quella della flessibilità normativa. L'art. 96, par. 1, lett. f), AI Act impone alla Commissione di elaborare orientamenti proprio in merito all'applicazione della definizione di sistema di IA.

elaborazione, decisione o apprendimento senza intervento umano diretto. In tale prospettiva, anche «i sistemi biologici od organici» rientrano nella categoria dei sistemi automatizzati purché abbiano delle capacità computazionali⁵⁰.

La definizione normativa ha un forte significato dommatico: ogni sistema di AI è anzitutto un sistema machine-based, ma non ogni sistema computazionale fondato su macchine integra un sistema di AI. Infatti, per essere qualificato come sistema di AI, il machine-based system deve essere progettato per operare con livelli variabili di autonomia e possedere capacità inferenziali, ossia essere in grado, sulla base di obiettivi espliciti o impliciti, di desumere dagli input ricevuti come generare output — quali previsioni, contenuti, raccomandazioni o decisioni — suscettibili di influenzare ambienti fisici o virtuali.

Quanto ai «livelli di autonomia variabili», il considerando 12 chiarisce che i sistemi di AI «dispongono di un certo grado di autonomia di azione rispetto al *coinvolgimento umano* e di capacità di funzionare senza *l'intervento umano*»⁵¹. Dunque, il concetto di autonomia viene sviluppato in relazione al grado di coinvolgimento e di intervento umano nel funzionamento del sistema. Emerge così la tensione di fondo presente nell'interazione uomo-macchina: da un lato, la spinta verso un'automazione che possa disimpegnare l'uomo da determinate attività ritenute gravose, routinarie ecc., fino a raggiungere livelli tali che i sistemi sono capaci «di funzionare senza l'intervento umano»; dall'altro, l'esigenza di confinare questa capacità nell'ambito di una categoria tecnica, non giuridica. In questa prospettiva, la posta in gioco non è stabilire se la macchina sia o meno intelligente. Essa, sul piano giuridico, non ha autonomia ma è uno strumento sofisticato la cui capacità di azione deve restare compatibile col modello di tutela dei diritti fondamentali centrato sull'uomo⁵². In altri termini, da un lato si riconosce l'elevato grado di sofisticazione tecnica e funzionale dei sistemi di AI (autonomia, adattabilità e capacità inferenziali), dall'altro, si tende ad ascriverli comunque alla categoria generale dei sistemi automatizzati (machine based) che operano sempre sulla base di istruzioni, parametri e obiettivi determinati dall'uomo⁵³.

⁵⁰ Comunicazione della Commissione, *Orientamenti della Commissione sulla definizione di sistema di intelligenza artificiale stabilita dal regolamento (UE) 2024/1689 (regolamento sull'IA)*, Bruxelles, 29.7.2025 (C 5053 final), § 13.

⁵¹ Corsivo aggiunto.

⁵² L'autonomia si genera nell'interazione uomo-macchina e può esser intesa come espressione della razionalità umana *regolata* dalla tecnica.

⁵³ Cfr. O. Pollicino, *Regolazione e innovazione tecnologica*, cit., 159, ove l'Autore afferma che «il regolamento enfatizza la capacità dell'IA di operare indipendentemente dall'intervento umano, ma nella pratica il confine non è netto. Alcuni sistemi automatizzati avanzati

Altro elemento che fonda tale lettura è il richiamo all'AI come «un sistema automatizzato *progettato* per funzionare con livelli di *autonomia* variabili»⁵⁴, per cui questa autonomia è strumentale a ottenere un sistema in grado di operare senza che ogni singolo output sia anticipatamente determinato dal progettista. Tale autonomia, pur differenziandosi dal paradigma deterministico dell'algoritmo inteso come sequenza predeterminata di istruzioni da eseguire passo dopo passo («*a set of step-by-step instructions*»)⁵⁵, al pari di questo non consente di qualificare le decisioni prodotte come entità giuridicamente indipendenti, non consente di discostarsi da quegli «obiettivi espliciti o impliciti» che teleologicamente guidano il sistema progettato per inferire previsioni, contenuti, raccomandazioni o decisioni che possono influenzare il contesto in cui operano. In questa prospettiva, l'autonomia riconosciuta dall'AI Act non postula una cesura tra l'operare dell'uomo e quello della macchina. Anzi si fonda proprio su questa relazione uomo-macchina, ossia un ambiente in cui alla macchina è affidato un determinato «ruolo» che condiziona la formazione della decisione, quel ruolo che l'art. 86 considera come uno degli elementi centrali della spiegazione. In termini giuridici, l'autonomia tecnica così delineata può essere ricondotta a una forma sofisticata di eteronomia, anche laddove il comportamento della macchina sia difficilmente prevedibile.

In conclusione, pur se la definizione fornita dall'art. 3 AI Act potrà determinare difficoltà interpretative e applicative, di cui il legislatore europeo è ben consapevole tanto da prevedere lo strumento degli orientamenti sull'applicazione del regolamento (art. 96 AI Act), essa raccoglie una sfida che ha un forte rilievo costituzionale: racchiude un'opzione ermeneutica di fondo laddove, nel riconoscere l'autonomia tecnico-funzionale, non ne fa discendere un'autonomia decisionale in senso giuridico.

Tale ricostruzione è di assoluto rilievo sul piano delle garanzie nel campo delle decisioni automatizzate: se il GDPR costruisce una dicotomia normativa tra decisione esclusivamente automatizzata e decisione nella quale l'intervento umano è giuridicamente significativo, l'AI Act prescinde da tale dicotomia perché, per definizione, il sistema di AI è progettato per funzionare con livelli variabili di autonomia, misurati in relazione alla sua indipendenza dal coinvolgimento umano. Non si tratta, dunque, dell'alternativa tra macchina e uomo, ma di un continuum di indipendenza fun-

potrebbero essere regolati come IA, mentre strumenti con effettiva capacità di apprendimento potrebbero sfuggire alla disciplina».

⁵⁴ Art. 3 AI Act (corsivo aggiunto).

⁵⁵ Cfr. Central Digital and Data Office – Centre for Data Ethics and Innovation, *Algorithmic Transparency Standard*, ove l'algoritmo è definito come «*a set of step-by-step instructions*». Lo Standard è stato sviluppato nell'ambito degli impegni assunti dalla *National Data Strategy* del Governo britannico.

zionale della macchina dall'uomo, per cui il baricentro della garanzia della spiegazione, intesa come controllabilità del potere mediato dalla tecnica, si sposta sul *sistema decisionale* e sul ruolo concretamente svolto dall'AI in tale contesto, come vedremo meglio di seguito.

4. Il diritto alla spiegazione nel GDPR: dal dibattito dottrinale alla giurisprudenza europea

Il GDPR dedica una specifica disciplina alle decisioni automatizzate. Si tratta dell'art. 22, rubricato "Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione"⁵⁶. L'architrave della tutela si fonda sul concetto di "non esclusività" della decisione automatizzata: «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» (par. 1)⁵⁷.

L'interpretazione dell'art. 22 GDPR ha posto e continua a porre significative sfide sia in dottrina che in giurisprudenza. L'intreccio delle sue previsioni normative, articolate in quattro paragrafi, ha rafforzato l'impressione di trovarsi di fronte a una «*house of cards*»⁵⁸, la cui solidità e funzionalità dipendono dalla capacità dell'interprete di approntare un'interpretazione coerente di ciascun elemento. Il quadro normativo, peraltro, è interessato

⁵⁶ Art. 22 GDPR: «1. 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. 2. Il paragrafo 1 non si applica nel caso in cui la decisione: a) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento; b) sia autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato; c) si basi sul consenso esplicito dell'interessato. 3. Nei casi di cui al paragrafo 2, lettere a) e c), il titolare del trattamento attua misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, almeno il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione. 4. Le decisioni di cui al paragrafo 2 non si basano sulle categorie particolari di dati personali di cui all'articolo 9, paragrafo 1, a meno che non sia d'applicazione l'articolo 9, paragrafo 2, lettere a) o g), e non siano in vigore misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato».

⁵⁷ Secondo le *Linee guida sulle decisioni automatizzate individuali*, adottate il 3 ottobre 2017 dal Gruppo di lavoro Articolo 29 per la protezione dei dati, istituito dall'art. 29 della direttiva 95/46/EC (c.d. A29WP), anche un annuncio pubblicitario 'personalizzato' a una decisione agli effetti dell'art. 22.

⁵⁸ Cfr. I. Mendoza-L. A. Bygrave, *The Right Not to be Subject to Automated Decisions Based on Profiling*, in T.E. Synodinou-P. Jougoux-C. Markou-T. Prastitou (a cura di), *EU Internet Law. Regulation and Enforcement*, Cham, 2017, 82.

da un progetto di modifica che potrebbe cambiare significativamente la disciplina dell'art. 22 GDPR⁵⁹.

Preliminarmente, possiamo annotare che la tutela normativa ricorre in presenza di due condizioni: (i) il legame della decisione con il trattamento automatizzato, che deve esser qualificato da una connessione forte che viene individuata dall'espressione «unicamente»; (ii) l'efficacia della decisione, che può manifestarsi nella produzione di «effetti giuridici» ovvero nell'incidere «significativamente» sulla sfera giuridica della persona⁶⁰.

Già dalla lettera della norma possiamo trarre la sensazione di un ambito applicativo segnato sì da limiti pregnanti ma le cui sfumature danno adito a non poche questioni interpretative e che, in definitiva, consentono di indugiare sulla sua effettiva portata applicativa. Quanto alla decisione puramente automatizzata, la tutela non opera laddove l'automazione rivesta un ruolo secondario nella produzione dell'atto decisionale, eppure è proprio questa qualità che rischia di perdersi nella complessità della realtà. Indagare in quale misura il calcolo computazionale sia prevalente sull'elemento umano diviene l'indagine preliminare da condurre per verificare l'operatività della norma; diviene un'attività essenziale per poter invocare le tutele normative avverso le decisioni basate esclusivamente sul trattamento automatizzato e che potrebbero avere conseguenze significative per le proprie vite⁶¹. Eppure, in molti sistemi complessi, è difficile tracciare una linea netta, poiché le decisioni possono essere il risultato di interazioni complesse tra algoritmi e operatori umani⁶².

Anche il parametro dell'incidenza significativa della decisione sulla per-

⁵⁹ Commissione europea, *Proposta di regolamento «omnibus digitale»*, Bruxelles, 19 novembre 2025, COM(2025) 837 final, 2025/0360(COD).

⁶⁰ Cfr. CGUE, C-634/21, *SCHUEFA Holding AG* (2023), § 43, «L'applicabilità di tale disposizione è quindi soggetta a tre condizioni cumulative, vale a dire, in primo luogo, che deve esistere una “decisione”, in secondo luogo, che tale decisione deve essere “basata unicamente sul trattamento automatizzato, compresa la profilazione”, e, in terzo luogo, che essa deve produrre “effetti giuridici [riguardanti l'interessato]” o incidere “in modo analogo significativamente sulla sua persona”».

⁶¹ Il Gruppo di lavoro Articolo 29 per la protezione dei dati, *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679*, cit., 23, ha identificato almeno due elementi da considerarsi rappresentativi di un intervento umano: che il coinvolgimento umano sia opera di «una persona che dispone dell'autorità e della competenza per modificare la decisione» e che «tale persona dovrebbe prendere in considerazione tutti i dati pertinenti».

⁶² Sulla complessità dei sistemi decisionali integrati uomo-macchina, cfr. B. Wagner, *Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems*, in *Policy & Internet*, 1, 2019, 104 ss.; M. Almada, *Human Intervention in Automated Decision-Making: Toward the Construction of Contestable Systems*, in *International Conference on Artificial Intelligence and Law (ICAAIL)* New York, 2019, 2 ss.

sona non è di facile valutazione⁶³. Il punto cruciale sta proprio nella mancanza di criteri oggettivi attraverso cui valutare quando una decisione sia destinata ad incidere «significativamente» sulla persona dell'interessato⁶⁴. Ora, nel momento in cui ricorrono entrambe tali condizioni, opera il «diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato», diritto che peraltro non è di carattere assoluto, poiché può essere neutralizzato da basi giuridiche alternative, cioè da ipotesi di deroga molto ampie (necessità per la conclusione o l'esecuzione di un contratto; autorizzazione da parte del diritto dell'Unione o degli Stati membri; consenso esplicito dell'interessato - art. 22, par. 2). In letteratura si è molto dibattuto circa la natura giuridica dell'articolo 22, soprattutto perché la norma si esprime nei termini di diritto e non impone letteralmente un divieto. Per tali ragioni, la dottrina si è divisa tra chi individua nella norma un divieto di carattere generale (*prohibition*)⁶⁵ e chi, invece, un diritto di opposizione in capo all'interessato al trattamento (*right*)⁶⁶. La Corte di giustizia interpreta la disposizione come un «divieto di principio» che non necessita di essere fatto valere individualmente⁶⁷. Per quel che rileva nella presente indagine, possiamo assumere che l'espressione adoperata dal legislatore non sia ambigua né tale da generare incertezze interpretative. Al contrario, si ritiene che la scelta terminologica

⁶³ Sulle criticità interpretative legate alla nozione di decisione che «incida in modo analogo significativamente sulla persona», si veda C. Sarra, *ICT 2.0. Decisioni automatizzate e sistemi contestabili: la via antica alla vera innovazione*, in C. Sarra (a cura di), *Il mondo-dato. Saggi su datificazione e diritto*, Padova, 2019, 137.

⁶⁴ Un'indicazione viene fornita nel considerando 71, laddove si cita a titolo esemplificativo «il rifiuto automatico di una domanda di credito online o pratiche di assunzione elettronica senza interventi umani». Anche il Gruppo di lavoro Articolo 29 per la protezione dei dati, *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679*, cit., 24, è intervenuto sull'argomento, enucleando quali situazioni significative quelle che influenzano le condizioni finanziarie di una persona (ammissibilità al credito), ovvero l'accesso ai servizi sanitari, che negano un'opportunità di impiego o pongono tale persona in una posizione di notevole svantaggio; decisioni che influenzano l'accesso di una persona all'istruzione, ad esempio le ammissioni universitarie.

⁶⁵ G. Malgieri-G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, in *International Data Privacy Law*, 4, 2017, 243 ss. In giurisprudenza, v. CGUE, C-634/21, *SCHUEA Holding AG* (2023), § 53, «A tal riguardo, occorre rilevare che, come osservato dall'avvocato generale al paragrafo 31 di tali conclusioni, l'articolo 22, paragrafo 1, del RGPD conferisce all'interessato il "diritto" di non essere oggetto di una decisione fondata esclusivamente su un trattamento automatizzato, compresa la profilazione. Tale disposizione sancisce un divieto di principio la cui violazione non necessita di essere fatta valere individualmente da una tale persona».

⁶⁶ L.A. Bygrave, *Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision-Making*, in K. Yeung-M. Lodge (a cura di), *Algorithmic Regulation*, Oxford, 2019, 248 ss.

⁶⁷ Così CGUE, C-634/21, *SCHUEA Holding AG*, cit., § 53.

di qualificare la posizione giuridica dell'interessato come «diritto» – anziché qualificare la posizione in termini di «divieto» di porre in essere il trattamento unicamente automatizzato – si rivela coerente con l'approccio sotteso all'intero GDPR. Tale opzione lessicale discende dall'aver attribuito all'interessato una posizione attiva, volta a garantirgli strumenti effettivi di partecipazione e di controllo sul trattamento dei propri dati personali proprio in presenza di decisioni interamente automatizzate⁶⁸: «Nei casi di cui al paragrafo 2, lettere a) e c), il titolare del trattamento attua misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, almeno il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione» (art. 22, par. 3). Tale lettura non esclude che l'art. 22, par. 1, possa prevedere un divieto di principio⁶⁹.

Tra tali strumenti di partecipazione e controllo v'è il c.d. diritto alla spiegazione della decisione automatizzata (*right to explanation*), attorno alla cui natura si è sviluppato un intenso dibattito.

La dottrina che si è occupata del tema può essere schematicamente ricondotta a tre orientamenti principali. Un primo filone interpretativo enfatizza il peso dogmatico del mancato inserimento nella norma, diversamente che nel considerando 71⁷⁰, del diritto di ottenere una spiegazione della decisione *ex post*. Si porta in rilievo la differenza strutturale tra considerando e articoli: i primi hanno funzione eminentemente interpretativa e sistematica, fornendo criteri di orientamento per l'esegesi e l'applicazione delle norme; i secondi, al contrario, sono precetti giuridici vincolanti, dotati di forza prescrittiva e suscettibili di diretta applicazione. Alla luce di ciò, il considerando 71 non può essere interpretato come fonte autonoma di un diritto soggettivo alla spiegazione⁷¹. Allo stesso tempo, non si nega che gli obblighi informativi previsti dagli artt. 13 e 14 GDPR incidano sulla comprensibilità della decisione automatizzata ma si sostiene rientrino in

⁶⁸ Cfr. M. Brkan, *Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond*, in *International Journal of Law and Information Technology*, 2, 2019, 91 ss.; T. Zarsky, *The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making*, in *Science, Technology & Human Values*, 1, 2016, 118 ss.

⁶⁹ Cfr. CGUE, C-634/21, *SCHUEA Holding AG*, cit.

⁷⁰ Considerando 71 GDPR: «In ogni caso, tale trattamento [automatizzato] dovrebbe essere subordinato a garanzie adeguate, che dovrebbero comprendere la specifica informazione all'interessato e il diritto di ottenere l'intervento umano, di esprimere la propria opinione, di ottenere una *spiegazione della decisione* conseguita dopo tale valutazione e di contestare la decisione» (corsivo aggiunto).

⁷¹ 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.

un diritto all'informazione *ex ante* circa la logica generale del trattamento, da intendersi in chiave di trasparenza complessiva del processo decisionale automatizzato⁷².

Un secondo filone tende a dare rilievo a un'interpretazione sistematica richiamando gli obblighi rafforzati prescritti in capo al titolare del trattamento nel caso di impiego di decisioni unicamente automatizzate, quelli inerenti alla previsione di "misure adeguate" (art. 22, par. 3, GDPR) nonché quelli inerenti alle informazioni supplementari (artt. 13 e 14 GDPR)⁷³. In particolare, quanto agli oneri informativi, il titolare del trattamento ha l'obbligo di fornire all'interessato informazioni relative all'«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»⁷⁴. Tale posizione può essere ascritta anche al Gruppo di lavoro "Articolo 29", che nelle sue linee guida, seppur abbia precisato che il titolare del trattamento non debba fornire «necessariamente una spiegazione complessa degli algoritmi utilizzati o la divulgazione dell'algoritmo completo», per assolvere all'obbligo di fornire informazioni significative sulla logica utilizzata deve fornire informazioni «sufficientemente complete affinché l'interessato possa comprendere i motivi alla base della decisione»⁷⁵.

⁷² Cfr. F. Calisai, *I diritti dell'interessato*, in V. Cuffaro-R. D'Orazio-V. Ricciuto (a cura di), *I dati personali nel diritto europeo*, Torino, 2019, 336, ove l'Autore scrive che «Gli obblighi informativi sono senza dubbio da ritenersi funzionali all'esercizio dei diritti dell'interessato. Senza un'adeguata e trasparente circolazione e veicolazione delle informazioni sarebbe infatti precluso all'interessato stesso l'esercizio delle facoltà attribuitegli, che comprendono, come osservato, anche poteri di controllo e indirizzo».

⁷³ In tal senso, si veda C. Sarra, *Il diritto di contestazione delle decisioni automatizzate nel Gdpr*, in *Anuario Facultad de Derecho - Universidad de Alcalá* XII (2019), 57, ove l'Autore afferma che il diritto di spiegazione «debba ritenersi implicito come condizione di operabilità dei diritti riconosciuti, affinché essi non si svuotino di significato»; B. Goodman-S. Flaxman, *European Union regulations on algorithmic decision-making and a "right to explanation"*, in *AI Magazine*, 3, 2017, presentato per la prima volta nel 2016 all'ICML Workshop relativo al tema della Human Interpretability in Machine Learning (WHI 2016), New York, NY; M. Brkan, *Do algorithms rule the world?*, cit., 91 ss.

⁷⁴ Art. 13, par. 2, lett. f) ("Informazioni da fornire qualora i dati personali siano raccolti presso l'interessato") e art. 14, par. 2, lett. g) ("Informazioni da fornire qualora i dati personali non siano stati ottenuti presso l'interessato") del GDPR.

⁷⁵ Gruppo di lavoro Articolo 29 per la protezione dei dati, *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679*, adottate il 3 ottobre 2017 (versione emendata e adottata il 6 febbraio 2018) (WP 251 rev.01), 28. In un passaggio successivo, tuttavia, con riferimento al contenuto del diritto di accesso previsto dall'art. 15, par. 1, lett. h), GDPR, il Gruppo di lavoro sembra assumere una differente interpretazione, laddove scrive: «L'articolo 15, paragrafo 1, lettera h), afferma che il titolare del trattamento deve fornire all'interessato informazioni sulle conseguenze

Il terzo filone si inserisce nel dibattito con una posizione che tende ad attribuire maggior peso al considerando del regolamento. Infatti, sebbene venga riconosciuto che i considerando abbiano valore interpretativo e non forza vincolante, è stato sostenuto che al considerando 71 - in relazione all'art. 22, par. 3, GDPR - possa esser riconosciuto «a supplementary normative role»⁷⁶. Sulla base di tale passaggio ermeneutico si è ipotizzata la configurazione di un *right to legibility*. Il contenuto di tale diritto sarebbe differente rispetto al diritto di spiegazione, perché per leggibilità si intende la capacità degli individui di comprendere non tanto i dettagli tecnici del funzionamento dell'algoritmo, quanto piuttosto la logica, il significato e le conseguenze previste di un processo decisionale algoritmico. Si tratterebbe, quindi, di una situazione giuridica «more 'proactive'», cioè personalizzata in base al livello di comprensione individuale e diretta a garantire una concreta comprensibilità della logica e delle conseguenze⁷⁷.

Sebbene quest'ultimo filone sia particolarmente interessante, soprattutto perché sposta il dibattito da una visione rigidamente dicotomica tra c'è e non c'è un diritto alla spiegazione verso una prospettiva più attenta alla comprensibilità effettiva dei processi algoritmici, si ritiene preferibile l'orientamento che riconosce l'esistenza del diritto alla spiegazione. In effetti, a voler prendere «molto sul serio la previsione del diritto di contestazione», esplicitamente riconosciuto all'art. 22, par. 3, GDPR, ne consegue il riconoscimento di un vero e proprio diritto alla spiegazione della decisione funzionale a quello⁷⁸.

previste del trattamento, piuttosto che una spiegazione di una *particolare* decisione» (30).

⁷⁶ Sul punto si veda G. Malgieri-G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, in *International Data Privacy Law*, cit., 254-255, ove gli Autori notano come «it has been argued that recitals are not legally binding. Therefore, data controllers do not have a real duty to explanation according to Article 22(3) and Recital 71 GDPR. However, although it is undoubted that recitals are interpretative tools, they sometimes also perform a supplementary normative role. [...] In its illustrative nature, Recital 71 can be properly considered a supplementary normative tool».

⁷⁷ Ivi, 250.

⁷⁸ C. Sarra, *Il diritto di contestazione delle decisioni automatizzate nel Gdpr*, cit., 63. Questa interpretazione sembra trovare conferma nella recente decisione della Corte di giustizia del 27 febbraio 2025 nella causa C-203/22 (*D&B*), secondo cui «lo scopo principale del diritto dell'interessato di ottenere le informazioni previste all'articolo 15, paragrafo 1, lettera h), del RGPD è quello di consentirgli di esercitare in modo efficace i diritti riconosciutigli dall'articolo 22, paragrafo 3, di tale regolamento, vale a dire quello di esprimere il suo punto di vista su tale decisione e quello di contestarla» (§ 55). La Corte, altresì, precisa che «dal considerando 71 del RGPD si evince che, l'interessato, qualora sia oggetto di una decisione adottata unicamente sulla base di un trattamento automatizzato e che incida significativamente sulla sua persona, deve avere il diritto di ottenere una spiegazione in merito a tale decisione. Come rilevato dall'avvocato generale al paragrafo 67 delle sue conclusioni, si deve quindi ritenere che l'articolo 15, paragrafo 1, lettera h),

Su questa posizione si è attestata anche la Corte di giustizia, caso *Dun & Bradstreet Austria*⁷⁹. La controversia traeva origine dal rifiuto di un operatore telefonico di stipulare un contratto sulla base di una valutazione automatizzata della solvibilità elaborata da Dun & Bradstreet. Chiamata a precisare la portata delle «informazioni significative sulla logica utilizzata» previste dall'art. 15, par. 1, lett. h), GDPR, la Corte ha affermato espressamente che tale disposizione attribuisce all'interessato «un vero e proprio diritto alla spiegazione sul funzionamento del meccanismo alla base di un processo decisionale automatizzato di cui tale interessato è stato oggetto e sul risultato a cui detta decisione ha condotto»⁸⁰. La funzione di tale spiegazione è di consentirgli di esercitare effettivamente le garanzie previste dall'art. 22, par. 3. A tal fine, non è sufficiente né la semplice comunicazione «di una formula matematica complessa, come un algoritmo, né la descrizione dettagliata di tutte le fasi di un processo decisionale automatizzato»⁸¹, ma la descrizione, in forma concisa, trasparente, intelligibile e facilmente accessibile, «della procedura e dei principi concretamente applicati per utilizzare, con mezzi automatizzati, i dati personali dell'interessato al fine di ottenerne un risultato specifico, come un profilo di solvibilità»⁸².

Nel solco di quest'ultimo filone interpretativo, si può proporre una lettura ulteriore, coerente con la sua impostazione funzionale. Le garanzie previste dall'art. 22, par. 3, GDPR - il diritto a ottenere l'intervento umano, la possibilità di esprimere la propria opinione e la facoltà di contestare la decisione automatizzata - sono espressione di una dimensione partecipativa, all'interno della quale il nucleo essenziale è rappresentato dal riconoscimento di un vero e proprio *diritto al contraddittorio*. In altre parole, il diritto di intervento umano, di espressione e di contestazione, unitamente all'accesso a informazioni significative sulla logica e sugli effetti del trattamento (art. 13, par. 2, lett. f), art. 14, par. 2, lett. g), art. 15, lett. h), GDPR), si configurano come meccanismi di effettiva partecipazione e dialettica procedimentale, cosicché il diritto al contraddittorio rappresenta la sintesi funzionale di tali prerogative partecipative e si configura come posizione

del RGPD offra all'interessato un vero e proprio diritto alla spiegazione» (§ 57).

⁷⁹ CGUE, C-203/22, *Dun & Bradstreet Austria* (2025).

⁸⁰ Ivi, § 57.

⁸¹ Ivi, § 59.

⁸² Ivi, § 58. Uno degli elementi rilevanti della sentenza riguarda la relazione tra il diritto alla spiegazione e il segreto commerciale. Qualora le informazioni necessarie alla spiegazione contengano segreti commerciali o dati di terzi, il titolare del trattamento «è tenuto a comunicare tali informazioni asseritamente protette all'autorità di controllo o al giudice competente, cui spetta ponderare i diritti e gli interessi in gioco al fine di determinare la portata del diritto di accesso dell'interessato previsto all'articolo 15 del RGPD» (§ 76).

giuridica complessa che le ricomprende tutte, in modo unitario⁸³.

Come anticipato, il quadro normativo è interessato dal processo di revisione avviato con la proposta di Digital Omnibus⁸⁴. La Commissione propone, tra l'altro, di riformulare l'art. 22, par. 1, passando dalla previsione del «diritto di non essere sottoposto» alla decisione esclusivamente automatizzata all'individuazione diretta delle condizioni nelle quali tale decisione è consentita⁸⁵. La proposta potrebbe ampliare, in concreto, la possibilità di ricorrere alla decisione esclusivamente automatizzata, soprattutto nell'ambito contrattuale. In tale prospettiva, risulterebbe ulteriormente accentuata la funzione del diritto alla spiegazione quale strumento per la controllabilità della decisione.

Allo stato attuale, il quadro regolatorio risulta intriso di diritti e divieti, deroghe e prerogative, ma il punto cruciale è il rischio che le garanzie restino affermate su un piano meramente formale e che siano prive di effettiva incidenza sul piano sostanziale, ove la complessità dei sistemi algoritmici domina la scena. Proprio l'avvento dei sistemi dotati di autonomia, a volte intrinsecamente opachi o imperscrutabili per ragioni di privacy, espone il diritto al rischio di inadeguatezza: obblighi stringenti di spiegazione o di disclosure possono risultare, in concreto, inesigibili e tradursi di fatto in un divieto d'uso di tali tecnologie. Ora, se il diritto alla spiegazione è strumento di tutela della persona e garanzia di effettività dello Stato di diritto, allora è proprio su questo terreno che il costituzionalismo digitale è chiamato a giocare una partita difficile⁸⁶.

⁸³ Cfr. I. Mendoza-L.A. Bygrave, *The Right Not to be Subject*, cit.

⁸⁴ Commissione europea, *Proposta Digital Omnibus*, COM(2025) 837 final, cit.

⁸⁵ Cfr. European Data Protection Board (EDPB) – European Data Protection Supervisor (EDPS), *Joint Opinion 2/2026 on the Proposal for a Regulation as regards the simplification of the digital legislative framework (Digital Omnibus)*, adottata il 10 febbraio 2026. Il parere congiunto suggerisce di mantenere il divieto come principio, per evitare di insinuare che il processo decisionale automatizzato sia in linea di principio consentito ogniqualvolta vi sia un contratto, indipendentemente dal fatto che sia “necessario”.

⁸⁶ Come osserva A. Apostoli, *Per un costituzionalismo critico e per dei costituzionalisti critici*, in A. Apostoli-G. Azzariti (a cura di), *Diritto e politica. Per un costituzionalismo critico. Collana Costituzionalismo.it. Quaderno*, Napoli, 2025, la nozione di *costituzionalismo* è da intendersi «nella sua accezione classica di teoria dei limiti al potere, che va attualizzato e, in parte, riproblematizzato alla luce delle profonde trasformazioni che investono la forma e l'esercizio del potere nelle democrazie contemporanee. Non mi riferisco solo al potere politico-istituzionale in senso stretto, ma anche alle molteplici – e talora più pervasive – forme di potere economico, tecnologico, sovranazionale, che incidono profondamente sull'effettività dei diritti e sull'equilibrio tra gli organi e i livelli di governo» (3).

5. L'approccio sistematico dell'AI Act in termini di interpretabilità e spiegazione

L'AI Act affronta in modo esplicito la questione dell'autonomia dei moderni artefatti, e lo fa evidenziando la stretta relazione concettuale tra autonomia e automazione nonché intervenendo in modo significativo nel settore, particolarmente delicato, dell'autonomia delle decisioni. L'importanza di tale intervento regolatorio risiede anche nelle aspettative che gli sono state affidate⁸⁷. Si tratta, infatti, della prima disciplina organica e di ampia portata dell'intelligenza artificiale che, non senza critiche⁸⁸, si ritiene possa generare un effetto analogo a quello già osservato nell'ambito della protezione dei dati personali con il Regolamento generale sulla protezione dei dati, a proposito del c.d. Bruxelles Effect⁸⁹. Ne deriva l'esigenza di comprendere come il legislatore europeo abbia affrontato il tema della spiegabilità e della sua relazione con l'interpretabilità.

Per comprendere appieno la scelta di regolazione cui è giunto il legislatore europeo, pare opportuno muovere dalla dibattuta questione della spiegabilità (*explainability*) e interpretabilità (*interpretability*)⁹⁰ dei sistemi autonomi. Benché si tratti di una questione prevalentemente tecnica, essa assume

⁸⁷ Considerando 8 AI Act: «il presente regolamento contribuisce all'obiettivo di promuovere l'approccio antropocentrico europeo all'IA ed essere un *leader mondiale* nello sviluppo di un'IA sicura, affidabile ed etica» (corsivo aggiunto). Ancor più esplicita la Comunicazione della Commissione europea al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Promuovere un approccio europeo all'intelligenza artificiale*, 21.4.2021 COM(2021) 205 final, 5, «La normativa in materia di IA è agli albori e l'UE avrà molto da guadagnare se guiderà lo sviluppo di *norme globali nuove* e ambiziose e di iniziative di normazione e quadri di cooperazione internazionali in materia di IA» (corsivo aggiunto).

⁸⁸ M. Almada-A. Radu, *The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy*, in *German Law Journal*, 4, 2024, 646–663; B. Crum, *Brussels effect or experimentalism? The EU AI Act and global standard-setting*, in *Internet Policy Review*, 3, 2025; A. Engler, *The EU AI Act Will Have Global Impact, but a Limited Brussels Effect*, Brookings.edu (8 giugno 2022), (ultima visita 21.02.2026); C. Siegmann-M. Anderljung, *The Brussels Effect and Artificial Intelligence: How EU Regulation Will Impact the Global AI Market*, Technical Report, Centre for the Governance of AI, Agosto 22, 78.

⁸⁹ Così A. Bradford, *The Brussels effect: How the European Union rules the world*, New York, 2020. Il c.d. Brussels Effect indica la capacità dell'Unione europea di influenzare il “diritto vivente” di altri Paesi, posto che le imprese estere interessate ad entrare nell'attraente mercato europeo devono adeguarsi agli standard UE, standard che una volta adottati finiscono per essere applicati anche al di fuori dei confini europei e anche da parte di altre giurisdizioni.

⁹⁰ Sull'esistenza di un diritto all'interpretabilità tecnica, v. C. Gallese, *The AI Act proposal: a new right to technical interpretability?*, Extended Abstract, Information Society Law Center (ISLC), 20–22 February 2023, Online Event, disponibile su [arXiv](https://arxiv.org/abs/2302.00000) (ultimo accesso: 21 febbraio 2026).

rilevanza per uno studio reale dei problemi giuridici e per giustificare l'«arbitrio del legislatore»⁹¹, se l'obiettivo è elaborare regole volte a garantire la «protezione della salute, della sicurezza e dei diritti fondamentali [...] nonché promuovere l'innovazione»⁹².

In effetti, in letteratura i concetti di *explainability* e *interpretability* vengono spesso adoperati in modo indifferenziato, generando una certa confusione terminologica e concettuale⁹³. A voler sintetizzare le posizioni espresse, possiamo considerare l'*interpretability* come la proprietà intrinseca di un modello di essere comprensibile nella sua struttura e nel suo funzionamento, mentre l'*explainability* riguarda quei metodi basati su modelli secondari costruiti per fornire una giustificazione *ex post* del funzionamento di sistemi originariamente opachi (black box). Sul piano tecnico, l'AI spiegabile (XAI) rimane un obiettivo sfuggente e complesso⁹⁴, nel quale anche i metodi più all'avanguardia spesso forniscono spiegazioni errate, fuorvianti e incomplete⁹⁵.

A fronte di questo stadio delle conoscenze e dello sviluppo tecnologico, il regolamento ha adottato una prospettiva che muove in due direzioni: da un lato, introduce specifici obblighi di progettazione orientati alla comprensibilità operativa dei sistemi di AI ad alto rischio (art. 13 AI Act)⁹⁶, imponendo che essi siano corredati da «istruzioni per l'uso»⁹⁷ e informazioni

⁹¹ Si vuol richiamare la «dottrina realistica» di Tommaso D'Aquino, in R.M. Pizzorni, *La filosofia del diritto secondo S. Tommaso d'Aquino*, Bologna, 2003, 30-32, il quale «pone al centro o alla base del fenomeno giuridico la cosa o la realtà; cioè qualcosa di oggettivo, e non di soggettivo, come sarebbe la facoltà e come potrebbe essere anche la legge o la norma, ridotta all'arbitrio del legislatore» (30-32).

⁹² Considerando 1 AI Act.

⁹³ Per una posizione critica sulla confusione di parte della letteratura v. C. Rudin-C. Chen-Z. Chen-H. Huang-L. Semenova-C. Zhong, *Interpretable machine learning: Fundamental principles and 10 grand challenges*, in *Statistics Surveys*, 16, 2022, 1-85.

⁹⁴ Secondo C. Rudin, *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead*, in *Nature Machine Intelligence*, 5, 2019, 206-215, gli approcci di *Explainable ML* si basano su modelli secondari che spesso forniscono spiegazioni non affidabili o comunque fuorvianti.

⁹⁵ N.C. Chung-H. Chung-H. Lee-L. Brocki-H. Chung-G. Dyer, *False Sense of Security in Explainable Artificial Intelligence (XAI)*, in *International Joint Conference on Artificial Intelligence 2024 Workshop on AI Governance: Alignment, Morality, and Law*, Jeju, 2024.

⁹⁶ Articolo 13 AI Act (Trasparenza e fornitura di informazioni ai deployer): «1. I sistemi di IA ad alto rischio sono *progettati* e sviluppati in modo tale da garantire che il loro funzionamento sia sufficientemente trasparente da consentire ai deployer di *interpretare* l'output del sistema e utilizzarlo adeguatamente» (corsivo aggiunto).

⁹⁷ Considerando 72 AI Act, secondo cui le istruzioni per l'uso dovrebbero contenere «informazioni significative, complete, accessibili e comprensibili, tenendo conto delle esigenze e delle *conoscenze prevedibili* dei deployer destinatari» (corsivo aggiunto).

idonee a rendere intelligibili prestazioni, limiti e rischi⁹⁸; dall'altro, riconosce ai destinatari delle decisioni il diritto alla spiegazione (art. 86 AI Act), funzionale alla controllabilità del loro utilizzo. Dunque, la prospettiva giuridica non lascia spazio a una sovrapposizione concettuale tra explainability e interpretability e nemmeno attribuisce una posizione centrale alla prima. Se l'interpretabilità è una qualità del sistema che può e deve porsi come obiettivo della progettazione affinché il deployer possa comprendere il comportamento del sistema (non il funzionamento), e che acquisisce rilevanza nella relazione giuridica fornitore-deployer, la spiegabilità non è perseguita come caratteristica tecnica del sistema ma rappresenta l'oggetto di una specifica pretesa dell'interessato. Al di là del fenomeno puramente tecnico, si punta in modo diretto a responsabilizzare fornitori e deployer attraverso la configurazione di specifici obblighi informativi⁹⁹, nell'ambito di un più generale dovere di sorveglianza che vuole i sistemi di AI ad alto rischio progettati e sviluppati «in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui sono in uso» (art. 14 AI Act)¹⁰⁰.

Ai fini dell'analisi del diritto alla spiegazione, è utile ricostruire preliminarmente l'approccio sistematico seguito dal legislatore europeo, articolato almeno lungo due direttrici.

In primo luogo, l'intervento normativo si pone in linea di continuità con i rimedi già presenti nell'ordinamento, integrandoli con tutele ulteriori che si adattino alle trasformazioni tecnologiche. L'approccio sistematico si evince dalla clausola residuale: «Il presente articolo si applica solo nella misura in cui il diritto di cui al paragrafo 1 non sia altrimenti previsto dal diritto dell'Unione» (art. 86, par. 3, AI Act). La scelta normativa risponde a precise esigenze di sistema: da un lato, evita duplicazioni regolatorie che avrebbero prodotto obblighi ridondanti, incertezza applicativa e un au-

⁹⁸ Cfr. C. Casonato, *Unlocking the synergy: Artificial intelligence and (old and new) human rights*, in *BioLaw Journal*, 3, 2023, 233 ss., ove l'Autore pone in evidenza il rischio che il principio di supervisione umana, cui tende l'interpretabilità tecnica, resti confinato sul piano dell'affermazione di principio: per esercitarlo ci vorrebbe un professionista con specifiche competenze informatiche e che si assuma la responsabilità di contrastare l'output del sistema di AI, ci vorrebbe «un vero eroe» (238).

⁹⁹ Il regolamento (UE) 2026/1744, c.d. Digital Omnibus sull'AI, ha differito al 2 dicembre 2027 l'applicabilità delle sezioni 1-3 del capo III - quindi anche degli artt. 13 e 14 - per i sistemi ad alto rischio ex art. 6, par. 2 e Allegato III.

¹⁰⁰ L'art. 14, rubricato "Sorveglianza umana", prevede al par. 1 che «I sistemi di IA ad alto rischio sono progettati e sviluppati, anche con strumenti di interfaccia uomo-macchina adeguati, in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui sono in uso». Sui rapporti tra l'art. 14 AI Act e l'art. 22 GDPR, v. C. Sarra, *Artificial Intelligence in Decision-making: A Test of Consistency between the "EU AI Act" and the "General Data Protection Regulation"*, in *Athens Journal of Law*, 1, 2025, 45 ss.

mento non proporzionato degli oneri di compliance; dall'altro lato, evita di confondere strumenti con ambiti, funzioni e limiti distinti, come il GDPR, centrato sulla protezione dei dati personali, e l'AI Act, centrato sulla regolazione dei sistemi di AI secondo un approccio basato sul rischio.

L'effetto pratico è indubbiamente un ampliamento delle possibilità di tutela, attraverso rimedi posti in un rapporto di delimitazione ma anche di integrazione della disciplina¹⁰¹. L'effetto è un'espansione delle tutele in senso orizzontale: il GDPR presidia le decisioni esclusivamente automatizzate fondate sul trattamento di dati personali, mentre l'AI Act può intercettare le decisioni basate sull'impiego di sistemi di AI ad alto rischio che, pur non integrando i presupposti dell'art. 22 GDPR, risultano egualmente idonee a incidere in modo significativo sulla persona¹⁰². In questi termini, l'approccio sistematico registra l'effetto di una tutela complementare¹⁰³. Ne deriva non una duplicazione del medesimo diritto alla spiegazione, né una relazione che possa esaurirsi nello schema della *lex specialis* ma una continuità di protezione lungo fattispecie contigue che emergono con il diffondersi delle nuove tecnologie¹⁰⁴.

In secondo luogo, l'approccio di tipo sistematico conduce a configurare il diritto alla spiegazione con una natura ben precisa: l'art. 86 AI Act non attribuisce direttamente un potere di controllo ma un potere conoscitivo strumentale alla controllabilità della decisione. La mediazione algoritmica viene resa conoscibile affinché la decisione possa poi essere valutata e sottoposta ai rimedi disponibili nell'ordinamento¹⁰⁵. Si tratta, dunque, di un diritto di carattere informativo volto a integrarsi con istituti e rimedi di altri settori dell'ordinamento che attribuiscono direttamente poteri di

¹⁰¹ A. Häuselmann, *Déjà vu? An Analysis of Explanations Concerning Decision-Making Under the GDPR and the AI Act*, in *Journal of AI Law and Regulation*, 1, 2025, 37 ss.

¹⁰² La differenza principale è che, se il GDPR tutela la persona rispetto al trattamento dei suoi dati personali, l'AI Act tutela la persona perché esposta a un potere tecnologico che può incidere sulle libertà fondamentali, anche oltre la dimensione strettamente legata ai dati personali.

¹⁰³ Considerando 9 AI Act: «Le norme armonizzate stabilite nel presente regolamento dovrebbero applicarsi in tutti i settori e, in linea con il nuovo quadro legislativo, non dovrebbero pregiudicare il vigente diritto dell'Unione, in particolare in materia di protezione dei dati, tutela dei consumatori, diritti fondamentali, occupazione e protezione dei lavoratori e sicurezza dei prodotti, al quale il presente regolamento è *complementare*» (corsivo aggiunto).

¹⁰⁴ In termini di *lex specialis* si esprime L. Grozdanovski, *Non-discrimination law, the GDPR, the AI act and the - now withdrawn - AI liability directive proposal offering gateways to pre-trial knowledge of algorithmic discrimination*, in *AI Ethics*, 5, 2025, 5039–5062.

¹⁰⁵ E.H. van Beem-M. van Os, *Explainable throughout the AI lifecycle: the evolution of the explainability provisions in the EU AI Act*, in *International Review of Law, Computers & Technology*, 2026, 1–35, nel quale l'Autrice ricostruisce anche una specifica funzione delle spiegazioni come “contestation-enabling”.

contestazione. In questa prospettiva, il diritto alla spiegazione è autonomo quanto all'oggetto della pretesa informativa ma non configura una disciplina autosufficiente della contestazione della decisione¹⁰⁶.

Il confronto con il suo precedente più prossimo, il GDPR, rende evidente tale natura. In effetti, nel regolamento sulla protezione dei dati personali il diritto alla spiegazione è parte di una posizione giuridica complessa e contribuisce a qualificare il più ampio diritto al contraddittorio, fondato su oneri informativi rafforzati a carico del titolare del trattamento (artt. 13, par. 2, lett. f), e 14, par. 2, lett. g), GDPR)¹⁰⁷ e sulle tutele rafforzate minime del diritto all'intervento umano, del diritto di esprimere la propria opinione e del diritto di contestare la decisione (art. 22, par. 3, GDPR)¹⁰⁸. Nell'AI Act, invece, mancano garanzie partecipative equivalenti a quelle del GDPR, non vi sono autonomi poteri di interlocuzione, riesame e correzione della decisione. Ciò non esclude la funzione rimediale della spiegazione, vista peraltro la collocazione dell'art. 86 nella sezione dedicata ai «mezzi di ricorso», ma non trasforma la spiegazione in un rimedio direttamente correttivo della decisione. In questo senso depone anche il considerando 171 dell'AI Act, secondo cui la «spiegazione dovrebbe essere chiara e significativa e fornire una *base* su cui le persone interessate possano esercitare i loro diritti»¹⁰⁹. La formulazione normativa, quindi, sembra concepire la spiegazione precisamente come un antecedente informativo dell'esercizio di altre posizioni giuridiche¹¹⁰.

Tale funzione rimediale va letta anche alla luce della natura complementare del diritto alla spiegazione, destinato a integrarsi con ulteriori posizioni giuridiche e rimedi disponibili nell'ordinamento. La spiegazione può fornire elementi conoscitivi utili anche all'attivazione del reclamo previsto dall'art. 85 AI Act, un rimedio che opera sul distinto piano dell'*enforcement* e che non mira a consentire un contraddittorio individuale capace di determinare il riesame o la correzione della decisione.

In definitiva, l'art. 86 persegue un equilibrio complesso: da un canto, la questione tecnica della spiegabilità rappresenta ancora una domanda di

¹⁰⁶ Secondo L. Arnesen, *Comparing the right to information about automated decision-making to the new right to explanation*, in *International Data Privacy Law*, 2, 2026, ipag008, l'art. 86 AI Act rappresenterebbe «*a right without minimum safeguards*».

¹⁰⁷ Gli oneri informativi sono previsti anche in sede di diritto di accesso dall'art. 15, par. 1, lett. h), GDPR, relativamente al quale la CGUE, C-203/22, *Dun & Bradstreet c. Austria*, cit., ha chiarito che la spiegazione deve consentire all'interessato di comprendere e contestare la decisione e deve rendere intelligibili la procedura e i principi concretamente applicati.

¹⁰⁸ I diritti dell'art. 22, par. 3, GDPR, operano espressamente nelle ipotesi dell'art. 22, par. 2, lett. a) e c), GDPR.

¹⁰⁹ Corsivo aggiunto.

¹¹⁰ E. H. Van Beem, *Explainable throughout the AI lifecycle...*, cit.

ricerca aperta, per cui non la si può configurare come una caratteristica che debba necessariamente essere incorporata nei sistemi; dall'altro, per evitare qualsiasi suggestione o deriva deresponsabilizzante, la norma prende posizione sulla centralità della responsabilità umana e sulla necessaria funzione garantistica della spiegazione che, nella prospettiva della relazione uomo-macchina, riguarda la decisione «adottata dal deployer» e non il funzionamento del sistema automatizzato¹¹¹.

Ricostruito l'approccio sistematico del diritto europeo nel disciplinare il diritto alla spiegazione, resta da analizzare il contenuto di tale pretesa e la sua capacità di rendere effettivamente controllabile la decisione.

6. Contenuto e limiti del diritto alla spiegazione nell'AI Act

L'AI Act mira a consolidare il diritto alla spiegazione come strumento conoscitivo finalizzato alla controllabilità e, quindi, come condizione giuridica per l'esercizio effettivo dei diritti di difesa propri dello Stato di diritto. Tale esigenza risulta rafforzata in considerazione del fatto che i sistemi decisionali automatizzati possono essere dotati di livelli variabili di autonomia e di capacità inferenziali che possono rendere non immediatamente comprensibili i relativi output. Posto che l'art. 86 AI Act è autonomo quanto alla pretesa conoscitiva ma non autosufficiente quanto alla tutela, diviene necessario analizzare il contenuto del diritto alla spiegazione per valutare la capacità del rimedio di raggiungere il fine della controllabilità, soprattutto perché la disposizione fa riferimento a spiegazioni «chiare e significative», al «ruolo» del sistema e agli «elementi principali» della decisione senza individuare con maggiore precisione la densità informativa dovuta.

L'art. 86 dell'AI Act è rubricato “Diritto alla spiegazione dei singoli processi decisionali” ed è contenuto nella sezione 4 dedicata ai «mezzi di ricorso». Il par. 1 stabilisce che «Qualsiasi persona interessata oggetto di una decisione adottata dal deployer sulla base dell'output di un sistema di IA ad alto rischio ... e che produca effetti giuridici o in modo analogo incida significativamente su tale persona in un modo che essa ritenga avere un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi

¹¹¹ In disaccordo con quanto sostenuto da N.C. Chung-H. Chung-H. Lee-L. Brocki-H. Chung-G. Dyer, *False Sense of Security in Explainable Artificial Intelligence (XAI)*, cit., secondo cui «I regolatori e i politici sono dell'idea che la spiegabilità sia una questione di implementazione, piuttosto che un problema fondamentale di ricerca» (traduzione mia), l'AI Act punta a creare le condizioni giuridiche affinché l'innovazione possa procedere senza alibi di natura tecnicistica.

diritti fondamentali ha il diritto di ottenere dal deployer spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata»¹¹².

Dalla lettura della disposizione possiamo trarre alcuni punti chiave idonei a qualificare il diritto: (i) la rilevanza del *sistema decisionale* nella formazione della decisione, (ii) il diritto alla spiegazione sussiste soltanto con riferimento a sistemi di AI ad alto rischio che incidono sulla salute, sulla sicurezza o sui diritti fondamentali della persona, e (iii) il contenuto della spiegazione è delimitato dalla norma.

Quanto al primo aspetto, se il GDPR basa le sue garanzie sulla dicotomia uomo-macchina e richiede forme di tutela rafforzata in assenza dell'uomo (unicamente automatizzate), l'AI Act prescinde da tale dicotomia perché il sistema è progettato per funzionare in un continuum di indipendenza funzionale della macchina dall'uomo. La decisione è quindi «adottata dal *deployer* sulla base dell'output di un sistema di AI ad alto rischio»¹¹³ che è progettato per essere indipendente. In altri termini, sul piano giuridico assume rilevanza questa relazione uomo-macchina che qualifica un vero e proprio *sistema decisionale*, ossia un ambiente in cui alla macchina è assegnato un determinato ruolo nella formazione della decisione, quel ruolo che l'art. 86 prevede sia espressamente descritto affinché la spiegazione abbia un effettivo contenuto esplicativo.

Seguendo tale approccio, il contenuto della spiegazione non si concentra sulla determinazione computazionale ma verte «sul ruolo del sistema di AI nella procedura decisionale»¹¹⁴. Il regolamento non definisce espressamente che cosa debba intendersi per «ruolo del sistema di IA» e usa una nozione volutamente aperta che sappia contenere qualsiasi declinazione della relazione uomo-macchina. Come detto, non più la dicotomia uomo-macchina del GDPR, ove l'assenza dell'uomo abilita forme di tutela rafforzata, ma un sistema progettato per funzionare in un continuum di indipendenza funzionale della macchina dall'uomo. Tale ruolo può sostanziarsi in un contributo concreto nel classificare, prevedere, raccomandare o selezionare, in un contributo che interviene in una determinata fase della procedura e che può avere funzione consultiva, rilevante o dirimente ai fini della decisione. Un ruolo che finisce per definire lo spazio decisionale

¹¹² Art. 86 AI Act: «1. Qualsiasi persona interessata oggetto di una decisione adottata dal deployer sulla base dell'output di un sistema di IA ad alto rischio elencato nell'allegato III, ad eccezione dei sistemi elencati al punto 2 dello stesso, e che produca effetti giuridici o in modo analogo incida significativamente su tale persona in un modo che essa ritenga avere un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi diritti fondamentali ha il diritto di ottenere dal deployer spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata».

¹¹³ *Ibid.* (corsivo aggiunto).

¹¹⁴ *Ibid.*

integrato uomo-macchina, un ruolo che qualifica il *sistema decisionale*.

La concezione di un sistema decisionale, così inteso, non muta i termini dell'autorialità della decisione, posto che l'AI Act separa l'autonomia tecnico-funzionale del sistema dall'imputazione giuridica della decisione, che resta riferita all'unico autore umano (deployer), chiamato a risponderne secondo gli strumenti ordinari. Le decisioni automatizzate, quand'anche siano adottate con strumenti dotati di autonomia, rappresentano soltanto un raffinato prolungamento delle scelte e della razionalità umana che le ha progettate¹¹⁵. Si tratta pur sempre di una forma di manifestazione, seppur molto sofisticata ed efficiente, di eteronomia, in assenza di una qualche coscienza e comprensione etica che possa guidarne le scelte¹¹⁶. Per tale ragione, la decisione può essere «adottata dal deployer» soltanto, e la questione della *explainability* si traduce in un obbligo giuridico a carico del deployer per impedire qualsiasi alibi giuridico¹¹⁷. Detto altrimenti, il diritto affida ai fornitori (obbligati a progettare macchine interpretabili) e ai deployer dei sistemi di AI (obbligati a fornire spiegazioni delle proprie decisioni) la responsabilità di tradurre la trasformazione epistemologica in atto.

Quanto al secondo aspetto, relativo all'ambito applicativo del rimedio, occorre osservare che il diritto alla spiegazione non presenta una fisionomia di carattere generale. Infatti, l'art. 86 risulta molto circoscritto, in particolare ai sistemi ad alto rischio dell'Allegato III, escluso il punto 2, e alle ipotesi in cui la decisione produca effetti giuridici o in modo analogo incida significativamente su tali persone in un modo che esse ritengano avere un impatto negativo sulla loro salute, sicurezza o sui loro diritti fondamentali¹¹⁸. Tale ristretto ambito applicativo potrebbe farsi discendere dall'adozione del modello regolatorio *risk-based*: all'aumentare dei rischi, le prescri-

¹¹⁵ Cfr. L. Floridi, *AI as Agency without Intelligence: On Artificial Intelligence as a New Form of Artificial Agency and the Multiple Realisability of Agency Thesis*, in *Philosophy & Technology*, 1, 2025, art. 30, 18, il quale scrive che l'agenzia artificiale «is a computational, goal-driven form of agency defined by human purposes».

¹¹⁶ S. Chiodo, *Human autonomy, technological automation (and reverse)*, in *AI & Society*, 37, 2022, 39 ss., in cui l'Autrice esplora il passaggio dalla definizione delle tecnologie in termini di automazione a quella delle tecnologie in termini di autonomia come un cambiamento che «may mean that we are using technologies, and in particular emerging algorithmic technologies, as scapegoats that bear responsibility for us by making decisions for us».

¹¹⁷ In posizione critica rispetto all'attribuzione di autonomia ai sistemi di IA, per via del rischio di creare «lacune di responsabilità» in cui né gli esseri umani né le macchine possono essere ritenuti responsabili, sono F. Santoni de Sio-G. Mecacci, *Four responsibility gaps with artificial intelligence: Why they matter and how to address them*, in *Philosophy & Technology*, 4, 2021, 1057 ss.; A. Matthias, *The responsibility gap: Ascribing responsibility for the actions of learning automata*, in *Ethics and Information Technology*, 3, 2004, 175 ss.

¹¹⁸ M. Kaminski-G. Malgieri, *The Right to Explanation in the AI Act*, in *U of Colorado Law Legal Studies Research Paper*, 25-9, 2025, i quali osservano come l'art. 86 non sia corredato da autonomi diritti di contestazione dell'esito.

zioni di obblighi e diritti divengono più stringenti. Ciò non significa che la limitazione del diritto ai soli sistemi ad alto rischio sia una conseguenza automatica dell'adozione di un siffatto sistema ma che un'estensione generalizzata del diritto alla spiegazione ad altri o a tutti i sistemi di AI avrebbe comportato oneri e costi di compliance non coerenti con la proporzionalità tipica dei sistemi risk-based, col rischio di frenare l'innovazione e la diffusione tecnologica¹¹⁹. In tale prospettiva, l'art. 86 non configura un diritto generale alla spiegazione algoritmica, ma un rimedio calibrato sulle sole ipotesi in cui il potere decisionale automatizzato è idoneo a incidere più intensamente sulla sfera giuridica dell'individuo¹²⁰.

Infine, quanto al contenuto della prestazione informativa, oltre che al ruolo, il riferimento è ai «principali elementi della decisione adottata», cioè a quegli elementi che hanno influito sulla decisione. In questa prospettiva, la spiegazione non è concepita quale disclosure integrale della logica algoritmica né come accesso pieno al codice sorgente o ai modelli matematici sottostanti¹²¹. Essa si configura piuttosto come una spiegazione orientata alla comprensione del perché di una determinata decisione, sulla scia degli approdi cui si è giunti con riferimento al diritto alla spiegazione ex art. 22 GDPR. La sentenza *Dun & Bradstreet* individua il contenuto della spiegazione muovendo dalla sua funzione rispetto ai successivi poteri dell'interessato: comprendere le ragioni della decisione è necessario per poter esprimere il proprio punto di vista e contestarla¹²². Questa decisione può offrire un parametro interpretativo per concretizzare le spiegazioni chiare e significative sui «principali elementi della decisione adottata» richieste dall'art. 86, senza che ciò consenta tuttavia di sovrapporre le due discipli-

¹¹⁹ Cfr. F. Palmiotto, *The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the Legislative Process and the Future of the Regulation*, in *European Journal of Risk Regulation*, 2, 2025, 770–793, secondo cui il compromesso politico che ha caratterizzato il processo legislativo pone importanti sfide circa l'effettiva tutela dei diritti fondamentali nell'ambito dei sistemi di AI e soltanto in sede di applicazione e interpretazione giudiziaria della legge potremo verificare la reale tenuta del quadro regolatorio.

¹²⁰ Per una posizione critica v. M. Carta, *Il Regolamento UE sull'Intelligenza Artificiale: alcune questioni aperte*, in *rivista.eurojus.it*, 3, 2024, in particolare nt. 25; L. Nannini-A. Balayn-A.L. Smith, *Explainability in AI Policies: A Critical Review of Communications, Reports, Regulations, and Standards in the EU, US, and UK*, in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*, New York, 2023, 1198 ss.

¹²¹ Cfr. L. Edwards-M. Veale, *Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For*, in *Duke Law & Technology Review*, 2017, 18 ss.

¹²² Cfr. S. Wachter-B. Mittelstadt-C. Russell, *Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR*, in *Harvard Journal of Law & Technology*, 2, 2018, 841 ss., per i quali, essendo la spiegazione funzionale al controllo (e alla contestazione) dell'esito, piuttosto che all'intelligibilità piena del modello, essa dovrebbe consentire di comprendere come la decisione avrebbe potuto essere diversa al variare di determinati fattori rilevanti.

ne. Ciò che rileva, dunque, ai fini dell'assolvimento dell'obbligo di rendere una spiegazione non è tanto il modo in cui l'algoritmo abbia operato quanto il perché, nel caso concreto, sia stata adottata la decisione in quei determinati termini e quale sia stato il contributo dell'AI.

In definitiva, al fine di consentire un livello di conoscenza idoneo alla controllabilità del potere decisionale, la persona interessata ha diritto a spiegazioni chiare e significative sul ruolo dell'AI e sui principali elementi della decisione adottata. In concreto, la capacità del rimedio di assicurare la controllabilità rappresenta la vera sfida per lo Stato di diritto: rendere sufficientemente determinato il contenuto della spiegazione affinché essa possa assolvere effettivamente alla propria funzione di rendere controllabile il potere decisionale mediato dalla tecnica¹²³. La reale effettività del diritto alla spiegazione dipenderà dunque dalla capacità delle autorità, dei giudici e della dottrina¹²⁴ di tradurre la formula legislativa in standard concreti, idonei a rendere la spiegazione uno strumento effettivo di controllabilità del potere algoritmico e di responsabilità per chi lo utilizza¹²⁵.

¹²³ Come ha da tempo evidenziato M. Luciani, *Costituzionalismo irenico e costituzionalismo polemico*, in *Giurisprudenza costituzionale*, 2006, 1663: «una comunità politica ha la responsabilità di decidere sulle questioni che la agitano attraverso i propri organi politicamente responsabili, non può sempre scaricare sul circuito della giurisdizione il fardello della soluzione dei problemi»; v., *ex multis*, R. Hirschl, *Towards Juristocracy: The Origins and Consequences of the New Constitutionalism*, Harvard University Press, 2004, 286.

¹²⁴ Come osserva R. Romboli, *Il ruolo del giudice in rapporto all'evoluzione del sistema delle fonti ed alla disciplina dell'ordinamento giudiziario*, in *Quaderni dell'Associazione per gli studi e le ricerche parlamentari*, Seminario 2005, Torino, 2006, 81, «La produzione del diritto ad opera del legislatore e del giudice avviene certamente con forme e modalità differenti ed entrambi i soggetti operano nell'ordinamento secondo differenti forme di responsabilità e di legittimazione democratica. Come scrive recentemente Alessandro Pizzorusso, l'atto legislativo è rivolto direttamente a produrre diritto ed impone una regola per il futuro, mentre l'atto giurisdizionale è rivolto a risolvere una controversia sulla base di norme previgenti, con l'effetto solo indiretto di produzione normativa e, mentre il primo è fondato sulla volontà politica, il secondo trova il suo fondamento nella forza della ragione ed è qualificabile come fonte culturale».

¹²⁵ Cfr. U. Nnawuchi-C. George, *A Grand Entrance Without a Blueprint: A Critical Analysis of the Right to Explanation in Article 86 of the EU AI Act*, in *Journal of AI Law and Regulation*, 4, 2024, 402 ss., i quali evidenziano come l'attuale formulazione dell'art. 86 presenti una mancanza di «specificity and operational clarity»; B. Fresz-E. Dubovitskaya-M.F. Huber, *Explainable AI for the EU Right to Explanation: A Systematic Review of the Law-XAI Translation Gap*, arXiv preprint, arXiv:2608.02699, 2026, i quali affermano che, senza ulteriori progressi, il diritto alla spiegazione rischia di rimanere una «formal obligation without a technically realizable path to compliance».

7. Conclusioni

L'AI Act, e in particolare l'art. 86, si colloca in un disegno normativo volto a consolidare il diritto alla spiegazione quale garanzia di controllabilità del potere mediato dalla tecnica. La risposta del diritto europeo è certamente meritevole di apprezzamento perché introduce limiti e garanzie individuali che, in un'ottica di integrazione e di sistema, mirano a irrobustire i presidi di tutela della persona nell'ambito di una più generale attesa di ricondurre le decisioni automatizzate entro il perimetro dello Stato di diritto.

Lo sviluppo e la diffusione di sistemi decisionali automatizzati, in particolare di quelli dotati di livelli variabili di autonomia e di capacità inferenziali che possono rendere non immediatamente comprensibili gli output prodotti, rafforzano l'esigenza di conoscenza finalizzata alla controllabilità, e l'AI Act mira a consolidare il diritto alla spiegazione come strumento conoscitivo. Ma, proprio laddove aumenta il bisogno di comprensione, il grado di determinatezza del contenuto della spiegazione potrebbe condizionarne il fine della controllabilità. Il riferimento a spiegazioni chiare e significative, al ruolo del sistema e agli elementi principali non individua con precisione la densità informativa della prestazione dovuta, e sarà cura dell'interprete concretizzare progressivamente lo standard della spiegazione¹²⁶.

A ben vedere, questo profilo di genericità non è peculiare della sola disposizione in esame. Nel precedente dibattito sviluppatosi attorno al diritto alla spiegazione previsto dal GDPR non sono mancati analoghi rilievi. Ciò suggerisce che non si tratta di una criticità dovuta all'impiego di sistemi autonomi ma che essa si presenta ogniquale volta una decisione è il portato, in differente misura, dell'elemento computazionale, il quale introduce inevitabilmente una tensione tra razionalità tecnica e razionalità giuridica. Nella prospettiva di aderenza ai valori dello Stato di diritto qui considerata, assume rilievo la natura complementare del diritto alla spiegazione. Concepito come un rimedio non autosufficiente in termini di contestazione, ma complementare alle tutele previste da altri settori dell'ordinamento, il diritto alla spiegazione accentua la rilevanza della questione della determinatezza del suo contenuto: se la spiegazione è la base conoscitiva per l'esercizio di ulteriori rimedi, la sufficienza della prestazione informativa condiziona la stessa effettività del controllo. Una norma che apre a dif-

¹²⁶ Come osserva S. Mangiameli, *I diritti costituzionali: dallo Stato ai processi di integrazione*, cit., 9, «la normazione è un prodotto del gruppo organizzato e i suoi destinatari sono perciò tutti i componenti del gruppo. Per quanto paradossale possa sembrare, l'ordinamento vive dell'interpretazione che tutti fanno della normazione: i privati nella determinazione delle relazioni intersubiettive; la pubblica amministrazione nel perseguimento dei suoi fini; i giudici nell'esercizio della funzione giurisprudenziale; a questi si può aggiungere la dottrina giuridica con il suo ruolo scientifico».

ferenti letture aumenta il margine di variabilità applicativa, entrando in tensione con il principio generale della certezza del diritto¹²⁷. Questo non esige soltanto che le norme di diritto siano chiare e precise ma, seguendo l'interpretazione della giurisprudenza della Corte di giustizia e della Corte EDU, che la loro applicazione sia prevedibile per i soggetti dell'ordinamento¹²⁸. Non si dimentica l'insegnamento per cui lo Stato costituzionale di diritto contrappone al modello di Stato legislativo (in cui al vertice del sistema delle fonti v'è la legge, e il giudice solo a quella è soggetto) un modello in cui l'adeguamento del diritto ai principi costituzionali avviene grazie al contributo di molteplici soggetti, anche «per via interpretativa anziché per via legislativa»¹²⁹. La variabilità circa il contenuto dell'obbligo informativo affida all'opera collettiva di autorità amministrative, giudici e dottrina il compito di concretizzare lo standard sostanziale della spiegazione¹³⁰, dando certezza alle relative norme. La variabilità si traduce comun-

¹²⁷ Sulla certezza del diritto v., *ex multis*, A.C. Jemolo, *Il diritto come lo vedono i non giuristi*, Discorso inaugurale dell'anno accademico 1931-1932 all'Istituto Superiore di Scienze Economiche e Commerciali di Bologna, Bologna, 1932, secondo il quale la certezza è intesa come prevedibilità delle conseguenze giuridiche dei comportamenti degli individui; F. Carnelutti, *La certezza del diritto*, in *Rivista di diritto processuale civile*, 1943, 81 ss., per il quale il comando caratterizzato da certezza costituiva «il grande vantaggio di prevedere le conseguenze della loro condotta»; *Id.*, *Teoria generale del diritto*, III ed., Roma, 1951, 37, ove l'Autore scrive: «Certezza: la cui origine cernere, onde viene certezza, vuol dire vedere, il bisogno di certezza non è altro, in fin dei conti, che bisogno di vedere, quanto più chiaro possibile, nel futuro».

¹²⁸ CGUE, C-156/21, *Ungheria c. Parlamento europeo e Consiglio dell'Unione europea* (2022), § 223. Anche la Corte EDU ha insistito sulle esigenze di chiarezza, accessibilità e prevedibilità della base giuridica, esigenze che attengono alla certezza del diritto e al principio di legalità; tra tutte, v. CEDU, *The Sunday Times c. Regno Unito* (n. 1) (1979); CEDU, *Baranowski c. Polonia*, ric. 28358/95 (2000).

¹²⁹ In tal senso, G. Pino, *La certezza del diritto e lo Stato costituzionale*, in *Diritto pubblico*, 2, 2018, 542, per il quale il modello dello Stato costituzionale è «...caratterizzato dall'equilibrio e dal controllo reciproco tra molteplici poteri, dotati di titoli di legittimazione diversi (legislatore, autorità amministrative indipendenti, giudici comuni, Corte costituzionale, Corti sovranazionali)» per cui «la produzione legislativa del diritto è quasi sempre solo il passo iniziale di un processo metabolico».

¹³⁰ Cfr. R. Bin, *Certezza del diritto e legalità costituzionale*, in *Specula Iuris*, 2, 2023, 253 ss., ove l'Autore scrive: «La certezza del diritto entra nell'orizzonte del diritto costituzionale come uno degli obiettivi che danno senso all'intero sistema giuridico [...] il diritto, nel suo insieme, serve alla certezza dei rapporti sociali». Sulla necessità di distinguere tra certezza come fatto e certezza come valore v. G. Pino, *La certezza del diritto e lo Stato costituzionale*, cit., 518, il quale assume che «la certezza sia un valore meritevole di essere perseguito dal diritto [...] che la meritevolezza di questo valore (la sua desiderabilità su un piano etico-politico) non sia intaccata dal fatto che il diritto sia più o meno incerto». Già G. Zagrebelsky, *Il diritto mite. Legge diritti giustizia*, Torino, 1992, 202-203, scriveva: «non solo è dubbio che la certezza come fissità dell'ordinamento in tutte le sue parti possa essere oggi un obiettivo realistico. È dubbio che sia anche desiderabile».

que in una minore prevedibilità, almeno fino a quando lo standard non sarà consolidato, allontanando nell'immediato l'ambizione dell'Unione europea di assumere un ruolo di leadership globale, non solo normativa, nello sviluppo di un'AI sicura, affidabile ed etica.

Se la determinatezza del diritto alla spiegazione condiziona la sua effettiva capacità di rendere controllabile il potere mediato dalla tecnica, allora è su questo terreno che il costituzionalismo è chiamato a confrontarsi in una sfida complessa, assicurando che l'innovazione tecnologica resti compatibile con le garanzie fondamentali della persona.

Abstract

La diffusione dei sistemi decisionali automatizzati pone in evidenza la tensione tra la razionalità dell'automazione e la razionalità giuridica. La prima spinge a realizzare strumenti in grado di operare in modo sempre più autonomo rispetto al coinvolgimento umano e secondo processi inferenziali la cui complessità può eccedere la comprensione analitica dell'uomo. La razionalità giuridica mira a compensare questo divario attraverso il diritto alla spiegazione, al fine di ricondurre il potere mediato dalla tecnica entro un modello di controllabilità giuridica.

Il presente lavoro analizza l'approccio sistematico del diritto europeo che, con l'AI Act, consolida il diritto alla spiegazione come strumento informativo ma lo configura essenzialmente come un diritto non autosufficiente quanto alla contestazione della decisione. Poiché il diritto alla spiegazione è concepito come un rimedio complementare alle tutele previste da altri settori dell'ordinamento, il grado di determinatezza del suo contenuto condiziona l'effettività del rimedio stesso.

La sfida per lo Stato di diritto consiste nel rendere la spiegazione sufficientemente determinata da consentire alla persona di comprendere il ruolo svolto dall'intelligenza artificiale e di attivare effettivamente gli strumenti di controllo disponibili nell'ordinamento.

The spread of automated decision-making systems brings to the fore the tension between the rationality of automation and legal rationality. The former drives the development of tools capable of operating with increasing autonomy from human involvement and through inferential processes whose complexity may exceed human analytical understanding. Legal rationality seeks to bridge this gap through the right to explanation, with the aim of bringing technologically mediated power within a framework of legal controllability.

This paper examines the systematic approach adopted by European law, which, through the AI Act, consolidates the right to explanation as an informational instrument while essentially framing it as a right that is not self-sufficient for the purposes of challenging a decision. Since the right to explanation is conceived as a remedy complementary to the safeguards provided for in other areas of the legal order, the degree of determinacy of its content directly affects the effectiveness of the remedy itself.

The challenge for the rule of law is to ensure that explanations are sufficiently determinate to enable individuals to understand the role played by artificial intelligence and effectively activate the mechanisms of control available within the legal order.

Keywords

artificial intelligence, right to explanation, interpretability, AI Act, autonomy.