ISSN 1826-3534

federalismi.it
RIVISTA DI DIRITTO PUBBLICO ITALIANO, COMPARATO, EUROPEO

OSSERVATORIO TRASPARENZA
21 APRILE 2021

La trasparenza sul banco di prova dei modelli algoritmici

di Germana Lo Sapio
Magistrato del TAR Campania

La trasparenza sul banco di prova dei modelli algoritmici*

di Germana Lo Sapio
Magistrato del TAR Campania

Abstract [It]: La trasparenza costituisce una pietra miliare nella delineazione delle regole cui deve conformarsi l'applicazione di sistemi algoritmici, anche di intelligenza artificiale, nell'attività amministrativa. Tale principio emerge sia dal quadro normativo dell'Unione Europea in materia di intelligenza artificiale, in cui si persegue l'obiettivo di favorire lo sviluppo di una intelligenza artificiale antropocentrica, sia dalla giurisprudenza amministrativa che si è occupata delle decisioni fondate su algoritmi. Per entrambe le prospettive emerge un concetto di trasparenza intesa come "spiegabilità". Partendo dall'esame di questo scenario e dalla considerazione che i sistemi algoritmici presentano diversi livelli di opacità, questo articola affronta le seguenti questioni: a. il contenuto del diritto alla spiegazione, b. la spiegabilità come standard della progettazione, c. la spiegabilità come barriera all'ingresso dei sistemi di intelligenza artificiale caratterizzati dalla black box.

Abstract [En]: Transparency is a key milestone when outlining the scope of the application of algorithmic systems, including artificial intelligence, in administrative activity/environment. This principle underpines both the European Union's regulatory framework in the field of artificial intelligence, which pursues the objective of promoting the development of an anthropocentric artificial intelligence, and the administrative jurisprudence that has dealt with decisions based on algorithms. For both perspectives, a concept of transparency meant as "explainability" emerges. Starting from this regulatory scenario and algorithmic systems having different levels of opacity, this article highlights the followings: a. the content of the right to explanation, b. how explainability can become a standard of design, c. how the dilemma between explainability and artificial intelligence systems characterized by the black box must be solved**.**

Parole chiave: trasparenza, sistemi algoritmici, intelligenza artificiale, Amministrazione 4.0, spiegabilità
Keywords: transparency, algorithmic systems, artificial intelligence, Administration 4.0, explainability

1. La trasparenza nella prospettiva europea dell'Amministrazione 4.0.

Il dibattito sulla dimensione giuridica dell'innovazione tecnologica e digitale applicata all'attività amministrativa [1]si alimenta, oltre che del formante giurisprudenziale da ultimo emerso con riguardo alla

* Articolo sottoposto a referaggio.

[1] Riecheggiando la formula di "Industria 4.0." o "Quarta Rivoluzione Industriale", in cui le innovazioni digitali, l'interconnessione e l'automazione stanno determinando un ripensamento dell'organizzazione, della qualità e dei contenuti del lavoro, dei processi di produzione, di nuovi modelli di business, è stata coniata quella di Amministrazione 4.0., nella quale un ruolo determinante sarà sempre più assunto dall'intelligenza artificiale: cfr. D.U. GALETTA-J.G. CORVALÁN, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in*

cd. decisione fondata su algoritmi[2], del quadro regolatorio sempre più articolato che si va consolidando, soprattutto nell'ordinamento eurounitario.

Nel tentativo di tenere il passo rispetto alla impressionante velocità di cambiamento dei sistemi tecnologici che toccano ogni profilo della vita quotidiana[3], l'interesse delle Istituzioni europee si è manifestato, almeno a partire dal 2017, con una produzione incessante di documenti, relativi soprattutto al rapporto tra le tecnologie fondate sull'intelligenza artificiale e i principi etici, i valori fondanti dell'Unione Europea e i diritti fondamentali.

Da ultimo, nel "*grande fervore d'opera che sta caratterizzando l'interesse dell'Unione verso l'intelligenza artificiale*"[4] appaiono determinanti, da un lato, , il "*Libro bianco sull'intelligenza artificiale Un approccio europeo all'eccellenza e alla fiducia*"[5], che è stato adottato nell'ambito di un pacchetto di proposte del 19 febbraio 2020,

atto, in www.federalismi.it, n. 3/2019, 6 febbraio 2019, pp. 1 ss. In un più recente saggio, l'Autrice osserva che "*se guardiamo all'attività delle Pubbliche Amministrazioni, è palese che l'incrocio fra Big Data e Intelligenza Artificiale ha un potenziale enorme. Attraverso l'uso di sistemi di Intelligenza Artificiale è infatti possibile non solo gestire meglio tutti i dati e le informazioni in possesso delle Pubbliche Amministrazioni, ma è anche possibile una relativamente facile automazione di parte del processo decisionale, applicando l'Intelligenza Artificiale a vaste aree di attività di routine, ripetitive e standardizzate. L'introduzione nelle attività di back-office delle Pubbliche Amministrazioni di sistemi di Intelligenza Artificiale potrebbe consentire di automatizzare attività istruttorie che attualmente richiedono molti (e dispendiosi) passaggi interni. Con un evidente snellimento ed accelerazione complessiva dell'attività amministrativa ed una conseguente maggiore efficienza della macchina burocratico-amministrativa*"(D.U. GALETTA, *Algoritmi, procedimento amministrativo e garanzie: brevi riflessioni, anche alla luce degli ultimi arresti giurisprudenziali in materia*", in Rivista Italiana di Diritto Pubblico Comunitario, 2020, n. 3, p. 501).

[2] Le questioni giuridiche derivanti dall'applicazione di algoritmi al servizio delle funzioni amministrative e nell'ambito del procedimento hanno ricevuto, tra il 2018 e il 2019, una diffusa attenzione per effetto del filone giurisprudenziale sulla procedura di mobilità dei docenti, nell'ambito dal piano straordinario di assunzioni di cui alla legge 107/2015. Le assegnazioni delle sedi erano state gestite mediante un algoritmo – elaborato da fornitori terzi - poi rivelatosi fallace nei risultati, poiché i docenti erano stati trasferiti in province più lontane da quella scelta con priorità in sede di partecipazione alla procedura, benché nelle province di elezione i posti fossero rimasti vacanti. Si segnalano, in particolare anche per i precedenti ivi citati, Cons. Stato Sez. VI, 13 dicembre 2019, n. 8472, con commento dai R. MATTERA, *Processo - decisioni algoritmiche. Il Consiglio di Stato fissa i limiti*, in *Nuova Giur. Civ., 2020*, 4, 809; A. MASCOLO, *Gli algoritmi amministrativi: la sfida della comprensibilità*, in *Giornale Dir. Amm.*, 2020, 3, 366; M. TIMO, *Algoritmo - Il procedimento di assunzione del personale scolastico al vaglio del Consiglio di Stato*, in *Giur. It.,* 2020, 5, 1190; nonché, Cons. Stato Sez. VI, 04 febbraio 2020, n. 881 con commento di A. G. OROFINO – G. GALLONE, *Intelligenza artificiale - l'intelligenza artificiale al servizio delle funzioni amministrative*: profili *problematici e spunti di riflessione, in Giur. It., 2020, 7,*

3 K. KELLY, cofondatore della rivista di cultura digitale Wired, individua come prima direttrice del cambiamento tecnologico proprio quella del suo incessante divenire, se si considera che le più importanti tecnologie nella vita tra trent'anni devono ancora essere inventate; tutte le nuove tecnologie richiedono continui aggiornamenti e il loro ciclo di invecchiamento riduce progressivamente, con conseguente contrazione del tempo necessario ad acquisirne padronanza prima che vengano sostituite, con l'effetto di rendere permanente lo stato di sentirsi principianti ed inesperti ("newbie", tradotto nell'edizione italiana come "niubbo"): *"Eterno niubbo è il nuovo standard per tutti, a prescindere dall'età o del livello di esperienza",* il che, secondo Kelly, *"dovrebbe mantenerci umili*" (K. KELLY*, L'inevitabile, Milano,* 2017, p. 19)

[4] A. ADINOLFI, *L'Unione europea dinnanzi allo sviluppo dell'intelligenza artificiale: la costruzione di uno schema di regolamentazione europeo tra mercato unico digitale e tutela dei diritti fondamentali,* in *(*a cura di S. DORIGO)*, Il ragionamento giuridico nell'era dell'intelligenza artificiale*, Pisa, 2020, p.17.

[5] 19.2.2020 COM(2020)65. Nel tentativo di offrire anche una nozione condivisa di intelligenza artificiale, quale base di partenza per poterne regolare i profili giuridici, la Commissione definisce l'IA come "*un insieme di tecnologie che combina dati, algoritmi e potenza di calcolo*", facendo emergere, anche sotto l'aspetto definitorio, la stretta correlazione tra la sempre più ampia disponibilità di dati (cd. Big Data) e lo sviluppo dei modelli di intelligenza artificiale ("*i progressi compiuti nell'ambito del calcolo e la crescente disponibilità di dati sono pertanto fattori determinanti per l'attuale crescita dell'IA*").

comprendenti anche la comunicazione quadro "*Plasmare il futuro digitale dell'Europa*"[6] e la comunicazione sulla "*Strategia europea per i dati*"[7]e, dall'altro, ancora più recentemente, la "*Risoluzione del Parlamento europeo del 20 gennaio 2021 Intelligenza artificiale: questioni relative all'interpretazione e applicazione del diritto internazionale'*[8].

L'approccio dell'Unione Europea nei confronti delle nuove tecnologie si muove su un binario.

Emerge innanzi tutto l'obiettivo di recuperare terreno rispetto anche agli altri Stati (soprattutto Stati Uniti e Cina), al fine di aumentare la capacità innovativa delle imprese e la competitività del mercato europeo e - in relazione all'uso di intelligenza artificiale nei rapporti tra cittadini e pubblica amministrazione - di cogliere le opportunità di miglioramento della vita e di semplificazione dei meccanismi decisionali che da esse possono derivare. Ma, nella consapevolezza della vis espansiva degli sviluppi tecnologici più avanzati e dei rischi ad essi connessi, in tutti i documenti dell'Unione Europea si afferma la necessità che lo sviluppo del "mercato digitale" e l'adozione dell'I.A. nel settore pubblico[9] siano coerenti e compatibili con i diritti fondamentali e i valori (talvolta indicati come "principi etici") dell'ordinamento giuridico europeo.

Viene così a configurarsi, lungo questa seconda direzione, una versione giuridico-sostenibile, secondo un approccio mutuato dalle strategie normative in materia ambientale, dello sviluppo tecnologico, fino ad elaborarsi la nozione di "intelligenza artificiale antropocentrica"[10], posta al servizio dell'uomo e sotto la

---

[6] 19.2.2020 COM(2020)67

[7] 19.2.2020 COM(2020)66

[8] P9_TA-PROV(2021)0009, con la quale il Parlamento rappresenta l'esigenza di disporre di un quadro giuridico europeo comune, con definizioni armonizzate e principi etici comuni, sull'intelligenza artificiale che comprenda anche il suo utilizzo a fini militari (campo che era invece rimasto escluso nel Libro Bianco della Commissione del 2020). Nella Risoluzione, viene così sollecitata una definizione più dettagliata di intelligenza artificiale, da intendersi come un "*sistema basato su software o integrato in dispositivi hardware che mostra un comportamento che simula l'intelligenza, tra l'altro raccogliendo e trattando dati, analizzando e interpretando il proprio ambiente e intraprendendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici*".L'approccio verso un uso dell'intelligenza artificiale compatibile con i valori dell'Unione enunciati ora nell'art. 2 TUE comporta l'esigenza che, nell'intera catena del valore nello sviluppo, nell'attuazione e negli impieghi delle innovazioni tecnologiche sia assicurato il rispetto dei "*valori umanistici intrinsecamente europei e universali*"; che vi siano "*criteri rigorosi per controllarne, tra l'altro, la sicurezza, la trasparenza, la rendicontabilità, la non discriminazione e la responsabilità sociale e ambientale";* che "*là dove l'intelligenza artificiale è utilizzata dalle autorità pubbliche siano garantite la spiegabilità degli algoritmi, la trasparenza e la sorveglianza regolamentare, e che siano effettuate valutazioni d'impatto prima del ricorso da parte di autorità statali a strumenti che utilizzano tecnologie di IA".*

[9] Una delle azioni proposte nel Libro Bianco è quella della promozione delle applicazioni di I.A. nei settori pubblici: "*È essenziale che le amministrazioni pubbliche, gli ospedali, i servizi di pubblica utilità e di trasporto, le autorità di vigilanza finanziaria e altri settori di interesse pubblico inizino rapidamente a utilizzare nelle loro attività prodotti e servizi che si basano sull'IA. Un'attenzione particolare sarà rivolta ai settori dell'assistenza sanitaria e dei trasporti, in cui la tecnologia è abbastanza matura da consentire una diffusione su vasta scala. La Commissione avvierà dialoghi settoriali aperti e trasparenti dando priorità agli operatori del servizio pubblico, delle amministrazioni rurali e dell'assistenza sanitaria, al fine di presentare un piano d'azione che faciliti lo sviluppo, la sperimentazione e l'adozione dell'IA. I dialoghi settoriali saranno utilizzati per elaborare uno specifico "Programma di adozione dell'IA", che sosterrà gli appalti pubblici di sistemi di IA e contribuirà a trasformare le procedure stesse degli appalti pubblici*", Libro Bianco cit., p. 9

[10] L'approccio era già stato indicato nella comunicazione del 25 aprile 2018 "Strategia per l'I.A" COM(2018) 237 final; il Consiglio europeo del giugno 2018 ha poi sollecitato la Commissione a collaborare con gli Stati membri per definire un piano coordinato in materia di intelligenza artificiale; è stato pertanto istituito, con la comunicazione del 7 dicembre 2018 un gruppo di esperti sull'intelligenza artificiale che nel 2019 ha elaborato "orientamenti" per una I.A. affidabile, elencando sette specifici requisiti (intervento e sorveglianza umani, robustezza tecnica e sicurezza, riservatezza e

sua costante sorveglianza, una intelligenza "affidabile" che per essere tale deve conformarsi a specifici requisiti fondamentali, tra i quali un rilievo determinante è assunto da quello della trasparenza.

La conformità dei sistemi di I.A. al principio di trasparenza è esaminata da due speculari punti di vista: nella prospettiva interna alle specifiche applicazioni, la trasparenza corrisponde alla "*tracciabilità*", ovvero alla necessaria registrazione e documentazione sia delle decisioni adottate "*dai sistemi*" che dell'intero "*processo che ha prodotto le decisioni, comprese una descrizione della raccolta e dell'etichettatura dei dati e una descrizione dell'algoritmo utilizzato*"; nella prospettiva esterna che guarda ai destinatari delle applicazioni di intelligenza artificiale essa è intesa come "*spiegabilità del processo decisionale degli algoritmi, adattata alle persone coinvolte*". Da questo angolo visuale, la Commissione sottolinea l'esigenza che "*dovrebbero anche essere disponibili spiegazioni sulla misura in cui un sistema di IA influenza e definisce il processo decisionale organizzativo, le scelte di progettazione del sistema e la logica alla base della sua diffusione, in modo da garantire la trasparenza non solo dei dati e dei sistemi, ma anche dei modelli di business E' importante comunicare opportunamente e in modo deguato al caso in esame, capacità e limiti del sistema di IA ai diversi portatori di interessi coinvolti. I sistemi di IA dovrebbero essere identificabili come tali, così che gli utenti sappiano che stanno interagendo con un sistema di IA e possano individuare le persone che ne sono responsabili*"[11].

2. Gli ostacoli alla effettiva attuazione del principio di trasparenza: l'opacità "variabile" dei sistemi algoritmici

L'attenzione sulla trasparenza, e la sua declinazione nel senso della spiegabilità, è accentuata, anche nei documenti dell'Unione Europea sopra citati, da una considerazione che riguarda le caratteristiche del fenomeno da regolare; caratteristica che emerge, con una particolare intensità, in molti modelli di intelligenza artificiale, ma che, sotto profili diversi, accomuna tutti i sistemi tecnologici incentrati sull'utilizzo di algoritmi: la loro opacità [12].

governance dei dati, trasparenza, diversità, non discriminazione ed equità, benessere sociale e ambientale, accountability), poi accolti nella comunicazione "Creare fiducia nell'intelligenza artificiale antropocentrica" COM(2019) 168 final.

[11] COM(2019) 168 final, p. 6.

[12] Il confronto tra competenze diverse, tecniche e giuridiche, che segnò gli albori dell'I.A., è ancora di più considerato oggi indispensabile anche per affrontare gli svariati profili giuridici che l'applicazione dei sistemi di I.A. comporta. Sul punto, cfr. A. G. OROFINO - G. GALLONE, *Intelligenza artificiale - l'intelligenza artificiale al servizio delle funzioni amministrative: profili problematici e spunti di riflessione*, Giur. It., 2020, 7, 1738, in cui si sottolinea l'esigenza di "*una consapevolezza tecnica sul funzionamento del mezzo, che normalmente non è appannaggio dei giuristi (…) il cui approfondimento richiede un proficuo confronto interdisciplinare tra operatori portatori di capacità ed esperienze variegate*". Quanto alla fuorviante similitudine tra l'intelligenza artificiale e quella umana, che è fonte di equivoci resistenti se non di pregiudizi per i "non addetti ai lavori", essa poggia sull'assunto, erroneo, che di quest'ultima siano conosciuti i meccanismi di funzionamento. La natura ingannevole della comparazione è efficacemente resa da Roger Penrose, matematico britannico e divulgatore scientifico, vincitore del Premio Nobel per la Fisica 2020, secondo cui, a differenza delle macchine che costituiscono un insieme di regole logiche coerenti, ma necessariamente incompleto poiché fondate su assiomi non derivabili automaticamente, la mente umana intuisce la "*verità*" dei concetti di fondo, è capace di comprendere ciò che conosce, e di farsi una rappresentazione del mondo perché la vera *"intelligenza richiede comprensione. E la comprensione richiede consapevolezza*"; capacità

In questa prospettiva, le prescrizioni in tema di trasparenza sono funzionali anche a verificare che siano rispettate tutte le altre regole, sia quelle eventualmente già vigenti e "*adeguatamente interpretate*" al mondo nuovo e variegato dei sistemi algoritmici (ad esempio in materia di protezione dei dati personali, o in materia di responsabilità per i danni da prodotti difettosi); sia quelle che, secondo una delle opzioni si strategia normativa ancora in discussione, possano in futuro regolare i diversi profili dell'economia digitale. In sostanza, proprio "*a causa della mancanza di trasparenza (opacità dell'IA) è difficile individuare e dimostrare eventuali violazioni delle disposizioni normative (comprese quelle che tutelano i diritti fondamentali), attribuire la responsabilità e soddisfare le condizioni per chiedere un risarcimento. Pertanto, al fine di garantire l'effettiva applicazione e il rispetto delle norme, può essere necessario adeguare o chiarire la legislazione in vigore in determinati settori, ad esempio in materia di responsabilità, come ulteriormente specificato nella relazione che accompagna il presente libro bianco*"[13]

Invero, partendo dal presupposto che i sistemi di intelligenza artificiale si pongono in rapporto di specie a genere rispetto ai sistemi algoritmici[14] può dirsi che: a) tutti gli algoritmi sono certamente caratterizzati dalla "opacità linguistica", dovuta alla circostanza che le istruzioni sono comunicate dal programmatore alla macchina in un linguaggio informatico, e non nel linguaggio naturale con cui sono espresse anche le regole giuridiche; b) molti di essi presentano un ulteriore livello di opacità, che può definirsi "giuridica", poiché l'innovazione tecnologica fondata su algoritmi, specie se acquistata dall'amministrazione sul mercato, può essere oggetto di diritti di proprietà intellettuale[15], o industriale riconosciuti in capo ai suoi "ideatori", connessi alla dinamiche di sfruttamento commerciale sulla segretezza dei codici sorgenti[16] (con

non presenti ad oggi in nessuna macchina. Cfr. Incontro fra Roger Penrose e Emanuele Severino tenutosi al Centro Congressi Cariplo di Milano il 12 maggio 2018 http://www.vita.it/it/article/2020/10/06/roger-penrose-lintelligenza-artificiale-non-esiste/156896/ visitato il 12 febbraio 2020.

[13] Libro Bianco, cit., p. 15

[14] Mentre l'algoritmo è una sequenza di istruzioni ordinate in modo preciso e chiaro al fine di trasformare dati di partenza (input) in un qualche risultato (output); nei sistemi di intelligenza artificiale, l'utilizzo di un insieme di algoritmi è supportato da una elevata capacità di calcolo ed è nutrito da una grande disponibilità di dati, eterogenei per forme e formati (Big Data). Appare peraltro utile distinguere, nell'area più generale degli algoritmi, da un lato, "*gli algoritmi deterministici*", in cui tutte le istruzioni, gli input e gli output attesi, nonché i passaggi necessari richiesti per produrre il risultato i criteri sono forniti ex ante dal programmatore, i quali presentano una logica lineare; dall'altro, "*gli algoritmi non deterministici*", in cui "*è la macchina stessa a darsi in tutto o in parte le stesse istruzioni, determinando i parametri che devono guidare la sua azione al fine di raggiungere il risultato*", con livelli di autonomia diversi a seconda della specifica tecnica impiegata. I sistemi di intelligenza artificiale rientrano in questa seconda categoria: cfr. per una ampia ricognizione, G. AVANZINI, decisioni amministrative e algoritmi informatici. Predeterminazione analisi predittiva e nuove forme di intellegibilità, Napoli, 2020, pp. 3-13

[15] cfr. T. FAELLI, *Le innovazioni in materia di tecnologia blockchain tra diritto dei brevetti e diritto d'autore*, in Dir. Industriale, 2020, 2, p.172; S. MAGELLI *Le nuove tecnologie nella giurisprudenza*, in Dir. Industriale, 2020, 2, p. 199; la questione del riconoscimento della brevettabilità dei sistemi algoritmici è al centro di un vivace dibattito, poiché, l'art. 45 comma 2 del D.Lgs. 10 febbraio 2005, n. 30 Codice della proprietà industriale, espressamente prevede che "*non sono considerate invenzioni*", suscettibili di essere oggetto di brevetto, tra l'altro "*i metodi matematici*" (lett.a) e i "*programmi per elaboratore*" (lett. b)

[16] Cfr. F. BRAVO, *Software di intelligenza artificiale e istituzione del registro per il deposito del codice sorgente*, in *Contratto e Impr.*, 2020, 4, 1412.

la conseguente necessaria esigenza di bilanciamento tra l'interesse conoscitivo e la tutela di tali diritti); c) per una specifica categoria di algoritmi, rientranti in una specifica sotto area dell'intelligenza artificiale, e in particolare, per quelli di cd. "*apprendimento automatico*" e di cd. "*apprendimento profondo*" (*machine learning* e *deep learning*), viene in rilievo una forma di opacità "strutturale" che deriva dallo stesso meccanismo di funzionamento del sistema, rimanendo oscuro ed impenetrabile anche per gli stessi programmatori comprendere come la macchina, partendo dai dati forniti, sia giunta ad un determinato risultato (è il fenomeno noto come "*black box*")[17].

Si tratta peraltro di una caratteristica che connota non tutte le tecnologie di I.A.; cosicché, anche in tale ambito, è opportuno distinguere, da un lato, i cd. "*sistemi esperti*", in cui la conoscenza di base (le informazioni di partenza, le regole da seguire e le procedure) è fornita alla macchina dall'uomo *ex ante* e in cui l'autonomia – che è poi la qualità che li connota come "intelligenti" – è costituita dalla capacità che essi hanno di dedurre nuovi fatti dai dati immessi o di elaborare quali informazioni aggiuntive sono necessarie; dall'altro, i più evoluti ed opachi sistemi di *machine learning* (e quelli tra essi di *deep learning*)[18] caratterizzati, almeno allo stato attuale dell'evoluzione scientifica, da una limitata interpretabilità.

Proprio in considerazione della mancanza di trasparenza dei sistemi più avanzati di intelligenza artificiale, che possono peraltro costituire il volano per il consolidamento del ruolo di leadership dell'Unione nel mercato digitale, la Commissione nel Libro Bianco del 2020 più volte citato sollecita pertanto la ricerca scientifica a muoversi nella direzione di "*creare collegamenti*" tra i due diversi approcci, combinando "*il ragionamento simbolico*" (ovvero i cd. sistemi cd. esperti) "*con le reti neurali profonde*" che sono caratterizzate dalla *black box*, al fine di "*rendere maggiormente spiegabili i risultati dell'IA*".

---

[17] "*Le caratteristiche specifiche di molte tecnologie di IA, tra cui l'opacità (effetto "scatola nera"), la complessità, l'imprevedibilità e un comportamento parzialmente autonomo, possono rendere difficile verificare il rispetto delle normative dell'UE in vigore volte a proteggere i diritti fondamentali e possono ostacolarne l'applicazione effettiva. Le autorità preposte all'applicazione della legge e le persone interessate potrebbero non disporre dei mezzi per verificare come sia stata presa una determinata decisione con il coinvolgimento di sistemi di IA e, di conseguenza, se sia stata rispettata la normativa pertinente. Le persone fisiche e giuridiche possono incontrare difficoltà nell'accesso effettivo alla giustizia in situazioni in cui tali decisioni possono avere ripercussioni negative su di loro",* Libro Bianco, cit., p. 13

[18] Si tratta dei sistemi che, potendo lavorare solo su una elevata quantità di dati, costituiscono la frontiera contemporanea dell'I.A., esplosa dagli anni '90 in poi, in concomitanza con la nascita e lo sviluppo del World Wide Web e della incessante disponibilità di dati e informazioni in formato digitali che esso ha offerto. Essi sono caratterizzati dalla capacità di auto-apprendere dall'esperienza accumulata nei data base di riferimento, ma necessitano di una idonea fase di addestramento su enormi quantità di dati. A differenza dei sistemi esperti, pertanto, non sono fondati su una logica simbolica, che reagisce in un modo predeterminato a precisi stimoli. Una chiara distinzione tra i diversi modelli algoritmici dell'intelligenza artificiale, e che utilizza la suddivisione tra sistemi simbolici e non simbolici per differenziare i sistemi esperti (simbolici) dagli altri, si rinviene in A. LONGO – G. SCORZA, *Intelligenza artificiale. L'impatto sulle nostre vite, diritti e libertà*, cit. pp. 38 e ss., in cui si evidenzia come i sistemi esperti presentano l'enorme vantaggio che "*si può facilmente spiegare perché è stata raggiunta una certa conclusione e quali sono stati i passaggi del ragionamento*", ma d'altro canto lo svantaggio di richiedere che tutte le conoscenze e le regole di funzionamento siano precodificate *ex ante*, con la conseguente loro utilizzabilità in concreto solo in ambiti delimitati e specializzati

3. La declinazione del principio di trasparenza nella giurisprudenza sulle "decisioni algoritmiche": il diritto alla spiegazione

Il doveroso ossequio al principio di trasparenza è stato indicato anche dalla giurisprudenza amministrativa, che ha affrontato i plurimi profili giuridici delle cd. decisioni algoritmiche nell'ambito[19], come il prioritario requisito di legittimità dell'utilizzo di sistemi algoritmici nel procedimento amministrativo; conclusione che certamente può estendersi anche al sotto insieme dei sistemi di intelligenza artificiale.

In sintesi, una volta risolta preliminarmente in senso positivo la questione del legittimo uso degli strumenti algoritmici nell'istruttoria procedimentale[20], la giurisprudenza amministrativa ha enucleato "*gli elementi di minima garanzia per ogni ipotesi di utilizzo di algoritmi in sede decisoria pubblica: a) la piena conoscibilità a monte del modulo utilizzato e dei criteri applicati; b) l'imputabilità della decisione all'organo titolare del potere, il quale deve poter svolgere la necessaria verifica di logicità e legittimità della scelta e degli esiti affidati all'algoritmo*".

Si tratta di requisiti di legittimità non posti su piani paralleli, ma legati da una precisa tassonomia, poiché, in mancanza della comprensione del meccanismo di funzionamento del sistema algoritmico anche da parte di chi lo utilizza nell'esercizio dell'attività decisoria (funzionario, responsabile del procedimento, e titolare dell'organo che ne valida gli approdi), la necessaria "*sorveglianza*" umana del processo decisionale rischia di essere un simulacro vuoto; diventando peraltro inesigibile anche il potere di "rivedere" le

---

[19] Cfr. nota 2.

[20] Secondo quanto osservato nella sentenza del Consiglio di Stato 8472/2020, "*la piena ammissibilità di tali strumenti risponde ai canoni di efficienza ed economicità dell'azione amministrativa (art. 1 l. 241/90), i quali, secondo il principio costituzionale di buon andamento dell'azione amministrativa (art. 97 Cost.), impongono all'amministrazione il conseguimento dei propri fini con il minor dispendio di mezzi e risorse e attraverso lo snellimento e l'accelerazione dell'iter procedimentale*" e sotto questo profilo è indubbio che la informatizzazione dell'attività amministrativa specie in attività seriali e ripetitive "*comporta infatti numerosi vantaggi quali, ad esempio, la notevole riduzione della tempistica procedimentale per operazioni meramente ripetitive e prive di discrezionalità, l'esclusione di interferenze dovute a negligenza (o peggio dolo) del funzionario (essere umano) e la conseguente maggior garanzia di imparzialità della decisione automatizzata".* Invero, va segnalato che, in data successiva all'adozione della predetta decisione, è stato novellato l'art. 3-bis della legge del procedimento amministrativo, 7 agosto 1990, n. 241, per effetto dell'art. 12, comma 1, lett. b), D.L. 16 luglio 2020, n. 76, convertito, con modificazioni, dalla L. 11 settembre 2020, n. 120., il quale, nella versione vigente dal 17 luglio 2020, prescrive che "*Per conseguire maggiore efficienza nella loro attività, le amministrazioni pubbliche agiscono mediante strumenti informatici e telematici, nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati*" segnando così un doppio cambio di passo: da un lato, la trasposizione nella sede naturale della legge sul procedimento amministrativo di un canone già evincibile dall'interpretazione sistematica dell'art. 12 del l'art. 12, comma 1, del CAD ("*Le Pubbliche Amministrazioni nell'organizzare autonomamente la propria attività utilizzano le tecnologie dell'informazione e della comunicazione per la realizzazione degli obiettivi di efficienza, efficacia, economicità, imparzialità, trasparenza, semplificazione e partecipazione*") e dell'art. 41 co. 1 "*Le Pubbliche Amministrazioni gestiscono i procedimenti amministrativi utilizzando le tecnologie dell'informazione e della comunicazione* "; dall'altro, la sostituzione del modello precedente, incentivante e meramente programmatico (l'art. 3-bis previgente prevedeva che "*Per conseguire maggiore efficienza nella loro attività, le Amministrazioni Pubbliche incentivano l'uso della telematica, nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati*") con un modello precettivo, tale da qualificare come doveroso il ricorso ad un metodo algoritmico se, tenuto conto dell'obiettivo assegnato all'azione amministrativa e della disciplina specifica, esso dia maggiori garanzie di efficienza rispetto all'opzione-zero.

determinazioni già assunte, ed eventualmente ritenute viziate per difetto di funzionamento dell'algoritmo, nelle forme dell'autotutela[21].

Viene così a configurarsi, anche nella prospettiva del diritto interno, una "*declinazione rafforzata del principio di trasparenza, che implica anche quello della piena conoscibilità di una regola espressa in un linguaggio differente da quello giuridico. Tale conoscibilità dell'algoritmo deve essere garantita in tutti gli aspetti: dai suoi autori al procedimento usato per la sua elaborazione, al meccanismo di decisione, comprensivo delle priorità assegnate nella procedura valutativa e decisionale e dei dati selezionati come rilevanti. Ciò al fine di poter verificare che i criteri, i presupposti e gli esiti del procedimento robotizzato siano conformi alle prescrizioni e alle finalità stabilite dalla legge o dalla stessa amministrazione a monte di tale procedimento e affinché siano chiare – e conseguentemente sindacabili – le modalità e le regole in base alle quali esso è stato impostato*"[22].

L'affermazione del principio di conoscibilità trova un addentellato normativo, ad avviso della giurisprudenza amministrativa che si riporta, nel diritto eurounitario, essendo riconosciuto in particolare, nel Regolamento UE n. 679/2016 del Parlamento Europeo e del Consiglio del 27 aprile 2016 sulla protezione dei dati personali (GDPR), il diritto a conoscere dell'esistenza di processi decisionali automatizzati che abbiano effetti giuridici sui destinatari e, anche nelle ipotesi in cui operino le deroghe al generale divieto di una decisione che sia esclusivamente fondata sul trattamento automatizzato di dati, quello a ricevere informazioni significative sulla logica utilizzata; principio a sua volta dedotto dall'art. 41 della Carta Europea dei Diritti Fondamentali ("*Right to a good administration*"). Anche in questo specifico contesto normativo, si discute dell'effettivo riconoscimento, in capo all'interessato, di un "*diritto alla spiegazione*" come ulteriore rispetto ai diritti di informazione ed accesso, di cui agli artt. 13, par. 2, lett. f, 14, par. 2, lett. g e 15, par. 1, lett. h. poichè, stando ad una interpretazione letterale del GDPR, non vi

[21] Il principio dell'imputabilità della decisione amministrativa, fondata su un algoritmo, all'organo titolare del potere evoca quello, visto dalla prospettiva del destinatario, di "non esclusività", ovvero la regola generale, prevista dall'art. 22 del GDPR, e già prima dall'art. 15 della Direttiva 95/46/CE, secondo cui una decisione che produca effetti giuridici che riguardano o incidano significativamente su una persona non deve essere "basata esclusivamente" su un processo automatizzato. A fronte di tale divieto generale, previsto nel par.1, la norma prevede però diverse eccezioni, previste dal successivo par. 2 che ne circoscrivono di molto l'efficacia prescrittiva. Il diritto ad una decisione non completamente automatizzata è infatti cedevole se: a) vi è il consenso esplicito dell'interessato; b) se il trattamento automatizzato è necessario per la conclusione o l'esecuzione di un contratto; c) se esso è autorizzato dal diritto dell'Unione o di uno Stato membro. In questo caso, e quindi nei casi in cui una decisione può fondarsi anche solo esclusivamente su trattamento "meccanizzato", resta comunque salvo il dovere del titolare del trattamento di predisporre le misure appropriate a garanzia dei diritti fondamentali e delle libertà dell'interessato, il quale ha comunque il diritto ad ottenere, nel caso concreto, l'intervento umano. Va anche sottolineato che, in relazione alle applicazioni di Intelligenza artificiale e quindi dinnanzi ad algoritmi complessi, basati sull'elaborazione di una mole di dati e considerati "affidabili", le perplessità sulla effettiva efficacia del principio di "imputabilità" della decisione finale all'uomo derivano dalla stessa *"travolgente forza pratica dell'algoritmo*" per cui, "*una volta introdotto un sistema automatico di decisione, all'interno del processo decisionale umano, il sistema automatico tende nel tempo a catturare la decisione stessa*" (A. SIMONCINI, *Diritto costituzionale e decisioni algoritmiche, in (a cura di S. DORIGO) Il ragionamento giuridico nell'era dell'intelligenza artificiale*, Pisa, 2020, p. 43; fenomeno molto preoccupante nelle applicazioni dell'intelligenza artificiale in materia di diagnostica medica.

[22] Cons. Stato, 8472/2019, cit.

sarebbe una perfetta sovrapposizione tra le "*misure appropriate*" che il titolare del trattamento deve assumere per tutelare i diritti le libertà e gli interessi legittimi dell'interessato ai sensi dell'art. 22 par.3 e le "*garanzie adeguate*" indicate nel 71° *Considerando*, secondo cui tali misure "*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*".

La declinazione della trasparenza come "diritto alla spiegazione", e quindi una interpretazione delle norme che tiene conto dell'evoluzione tecnologica, è invero imposta dalla caratteristica, sopra già evidenziata, che accomuna tutti i sistemi algoritmici, ovvero la loro "opacità linguistica", dietro il quale si nasconde il rischio non solo della non comprensibilità, ma anche della non corrispondenza tra ciò che prevede la regola giuridica e ciò che è stato "trasposto" nel linguaggio macchina. Invero, e' stato sul punto osservato che il passaggio dal linguaggio giuridico (che è a sua volta intriso di definizioni tecniche non pienamente sovrapponibili a quelle estrapolate dal linguaggio naturale) a quello informatico non è un'operazione scontata. Esso comporta di per sé "*un salto distruttivo e creativo*"[23] da parte del programmatore[24], il quale peraltro ha anche il compito di scegliere uno dei 2500 linguaggi di programmazione oggi rinvenibili[25].

Tenendo conto della asimmetria tra il linguaggio naturale e quello della macchina, e guardando all'effetto utile del principio di trasparenza, si impone pertanto, innanzi tutto, una "decodifica" di tale operazione di trasformazione linguistica.

In altri termini, il principio di trasparenza, che nel provvedimento amministrativo si articola nell'obbligo di motivazione, qualora sia utilizzato un sistema algoritmico, si traduce nell'obbligo di garantire al destinatario la capacità di comprendere autonomamente, senza il ricorso ad esperti esterni dotati di competenza informatica, il funzionamento, l'impatto, i dati utilizzati allo scopo, i criteri di elaborazione dei dati, in sintesi la "*ratio*" del processo decisionale[26].

---

[23] C. ACCOTO, *Il mondo dato. Cinque brevi lezioni di filosofia digitale*, Milano, 2017, p.21

[24] Cosicché, anche sul piano tecnico-informatico, è apparsa poco probante l'argomentazione difensiva dell'amministrazione, riportata nella sentenza del Consiglio di Stato n. 881/2020, secondo cui "*l'algoritmo è semplicemente il risultato della trasposizione matematica e della sua applicazione informatica delle direttive*",

[25] Una panoramica, facilmente fruibile, è evincibile in M. LAURELLI, *Dialoghi con una Intelligenza Artificiale*, Torino, 2020, pp. 34-42, dove si qualificano i linguaggi di programmazione a seconda del grado di vicinanza al linguaggio umano standard e quindi della loro leggibilità anche da parte degli stessi programmatori in linguaggi "*alti*" (come ad esempio il Fortran, "*alleggerito dalle tortuose sintassi binarie o numeriche: qualcosa di molto comprensibile per gli essere umani")* o *"bassi"* , vicini all'archetipo del codice binario, come il linguaggio Assembly. Peraltro, anche l'individuazione di un solo linguaggio informatico per una specifica applicazione non è così scontata come può sembrare. Nel caso dell'algoritmo utilizzato dal MIUR per la procedura di mobilità dei docenti del 2016 era emerso dalla consulenza tecnica che erano stati utilizzati almeno due linguaggi diversi, con evidenti ulteriori difficoltà interpretative.

[26] Una parte della dottrina propone pertanto "il concetto di legibility" volto a garantire la concreta comprensione del metodo e dei dati utilizzati; concetto che ha il merito di sintetizzare il "dilemma informazione/spiegazione", perché proiettato dalla parte di chi percepisce e non di chi eroga la spiegazionecfr. E. PELLECCHIA, *Profilazione e decisioni*

4. L'oggetto del diritto di spiegazione e la sua concreta operatività. Il modello della "*privacy by design*" esteso alla trasparenza delle decisioni algoritmiche

Riconosciuto il diritto alla spiegabilità, il problema pertanto assume ulteriori specifiche articolazioni, particolarmente complesse a fronte delle possibili applicazioni di intelligenza artificiale: a) cosa deve essere spiegato, ovvero quale livello di analiticità deve pretendersi, alla luce del quadro normativo vigente, per riempire di contenuto quel diritto; b) in che modo il canone della spiegabilità può essere veicolato nello statuto giuridico degli algoritmi posti a base della decisione; c) come si risolve il dilemma tra spiegabilità e la *black box*, che connota le applicazioni di intelligenza artificiale di *machine learning* e *deep learning*.

Sotto il primo profilo, appare persuasiva quella dottrina[27] che, sia pure con riguardo all'utilizzo di intelligenza artificiale nelle decisioni giurisdizionali, mette in luce la necessità di individuare esattamente quale sia il termine di paragone al fine di delineare il perimetro del dovere di spiegazione. In sostanza, prima di indicare cosa si pretende debba essere spiegato quando l'amministrazione pone a base della sua decisione un algoritmo, è necessario richiedersi quale è il contenuto necessario, ma anche sufficiente, della motivazione del provvedimento, la quale come è noto, deve indicare "*i presupposti di fatto e le ragioni giuridiche che hanno determinato la decisione dell'amministrazione in relazione alle risultanze dell'istruttoria*" che è volto "*a realizzare la conoscibilità, e quindi la trasparenza, dell'azione amministrativa*[28].

Se pertanto la motivazione è funzionale a ricostruire in modo puntuale l'iter logico seguito dall'amministrazione, per giungere, sulla base degli elementi acquisiti e valutati nell'istruttoria, ad una decisione; tale funzione segna anche il limite della sua esigibilità; pertanto, analogamente a come si esclude la necessità di un "fedele resoconto" del percorso mentale e cognitivo seguito dal funzionario che conduce l'istruttoria, secondo meccanismi che sfuggono anche alle scienze cognitive, appare quanto meno dubbio pretendere che la spiegazione del meccanismo di funzionamento di un algoritmo debba investire invece tutti ed ognuno dei passaggi logici seguiti dalla macchina. Resta in definitiva aperta la questione su quale sia il livello accettabile di interpretabilità o spiegabilità degli algoritmi, siano o meno essi riconducibili a sistemi di intelligenza artificiale.[29]

---

automatizzate al tempo della black box society: qualità dei dati e leggibilità dell'algoritmo nella cornice della responsible research And innovation, in Nuove Leggi Civ. Comm., 2018, 5, 1209.

[27] A. SANTOSUOSSO, I*ntelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto.* Milano, 2020, pp. 79-118. Sui profili processuali dell'uso di algoritmi, cfr. F. PATRONI GRIFFI, La decisione robotica e il giudice amministrativo, 28 agosto 2018, in www.giustizia-amministrativa.it

[28] Corte Cost., 5 novembre 2010, n. 310

[29] La discussione del limite della spiegabilità appassiona anche gli stessi esperti informatici, per i quali è doveroso chiedersi se il diritto alla spiegazione debba intendersi come quale diritto di "*capire come un algoritmo opera internamente*" o, diversamente, quale diritto a "c*apire quali fattori o dati siano stati decisivi nella produzione di un determinato risultato che ha avuto impatto su un individuo*" G. IOZZIA, *Trasparenza degli algoritmi. L'intelligenza artificiale deve essere spiegabile, ecco i progetti e le*

Quanto al secondo profilo, ovvero quello della concreta operatività del principio di trasparenza cui deve conformarsi l'uso di algoritmi nell'azione amministrativa, si va consolidando l'opinione che debba mutuarsi la logica, evincibile dall'art. 25 del GDPR [30], della cd. *privacy by design,* e collocare la regola della necessaria comprensibilità già alla fase anticipata della progettazione dei sistemi algoritmici. Il modello auspicato è un rovesciamento di prospettiva; la regola giuridica si integra nella progettazione del fenomeno da regolare, "gioca in anticipo" rispetto alla fase patologica della sua violazione, poiché conforma tecnicamente il suo oggetto dall'interno; in sostanza, mediante l'imposizione di vincoli di progettazione che rendano l'applicazione giuridicamente compatibile con i principi e diritti fondamentali, "*si potrebbe dire che una tecnologia, potenzialmente lesiva, viene privata della sua offensività per via tecnica. Il diritto governa questo movimento tecnico, ma fa un (mezzo) passo indietro rispetto alla tradizionale logico violazione/reazione*"[31].

Nell'ordinamento giuridico italiano, questo approccio trova anche un canale normativo. Ci si riferisce all'art. 13-bis del D.Lgs. 7 marzo 2005, n. 82 Codice dell'amministrazione digitale[32] (CAD), secondo cui i sistemi informatici e servizi digitali delle pubbliche amministrazioni devono essere progettati, realizzati e sviluppati "*in coerenza con gli obiettivi dell'agenda digitale italiana ed europea e nel rispetto del codice di condotta tecnologica*" (comma 1), specificando che "*il codice di condotta tecnologica disciplina le modalità di progettazione, sviluppo e implementazione dei progetti, sistemi e servizi digitali delle amministrazioni pubbliche, nel rispetto del principio di non discriminazione, dei diritti e delle libertà fondamentali delle persone e della disciplina in materia di perimetro nazionale di sicurezza cibernetica*" (comma 2)[33].

---

tecniche, in https://www.agendadigitale.eu/cultura-digitale/lintelligenza-artificiale-deve-essere-spiegabile-ecco-i-progetti-e-le-tecniche/ visitato il 12 gennaio 2021.

[30] Art. 25 GDPR "*Tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, come anche dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche costituiti dal trattamento, sia al momento di determinare i mezzi del trattamento sia all'atto del trattamento stesso il titolare del trattamento mette in atto misure tecniche e organizzative adeguate, quali la pseudonimizzazione, volte ad attuare in modo efficace i principi di protezione dei dati, quali la minimizzazione, e a integrare nel trattamento le necessarie garanzie al fine di soddisfare i requisiti del presente regolamento e tutelare i diritti degli interessati. Il titolare del trattamento mette in atto misure tecniche e organizzative adeguate per garantire che siano trattati, per impostazione predefinita, solo i dati personali necessari per ogni specifica finalità del trattamento. Tale obbligo vale per la quantità dei dati personali raccolti, la portata del trattamento, il periodo di conservazione e l'accessibilità. In particolare, dette misure garantiscono che, per impostazione predefinita, non siano resi accessibili dati personali a un numero indefinito di persone fisiche senza l'intervento della persona fisica*".

[31] A. SANTOSUSSO, *Intelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto,* cit., pag 189.

[32] inserito dall'art. 32, comma 1, D.L. 16 luglio 2020, n. 76, convertito, con modificazioni, dalla L. 11 settembre 2020, n. 120 (cd. decreto semplificazione 2020),

33 L'adozione del codice deontologico è di competenza del Capo dipartimento della struttura della Presidenza del Consiglio dei ministri competente per la trasformazione digitale, sentita l'AgID e il nucleo per la sicurezza cibernetica di cui all'articolo 12, comma 6, del decreto legislativo 18 maggio 2018, n. 65 e acquisito il parere della Conferenza unificata di cui all'articolo 8 del decreto legislativo 28 agosto 1997, n. 281. Con la finalità di garantire la effettiva attuazione della norma, si prevede anche che apetta all'AgID verificare il rispetto del codice di condotta tecnologica e che la sua violazione costituisce mancato raggiungimento di uno specifico risultato e di un rilevante obiettivo da parte dei dirigenti responsabili delle strutture competenti con le correlate conseguenze sulla valutazione delle performance individuali e sulla retribuzione.

Il Codice, che avrebbe dovuto essere adottato entro sessanta giorni dall'entrata in vigore della norma, verosimilmente si occuperà in maniera specifica degli standard minimi di qualità anche dei modelli algoritmici di intelligenza artificiale (esso risulta peraltro in fase di elaborazione al momento in cui si scrive) e può rappresentare la base normativo-regolatoria per rendere la trasparenza consustanziale alla struttura tecnica dei sistemi algoritmici utilizzabili dall'amministrazione. La norma primaria non richiama espressamente il principio di trasparenza, ma essendo la comprensibilità funzionale (anche) alla verificabilità del concreto operare dell'algoritmo, essa assume valore prioritario anche rispetto al principio – espressamente enunciato- di non discriminazione; poiché, come sottolineato dalla Commissione dell'Unione Europea, in mancanza di trasparenza, non è possibile neanche comprendere se siano state violate le legislazioni vigenti e, nello specifico, se effettivamente i risultati discriminatori siano dipesi dall'utilizzo di un set di dati che è già in partenza discriminatorio[34], o piuttosto da un erroneo funzionamento del programma.[35]

5. Il principio di trasparenza quale barriera all'ingresso dei sistemi algoritmici a "scatola nera"

Di fronte alla inderogabilità del principio di trasparenza, appare evidente come, per quanto avanzati e funzionali ad offrire opportunità di semplificazione, efficienza e aumento della qualità dei servizi erogati, i modelli di intelligenza artificiale connotati dalla cd. *black box*[36]non rispondano agli standard legali imposti affinché il loro utilizzo dia luogo a decisioni amministrative legittime. In realtà, l'esigenza di superare

[34] "*E' il principio noto tra i data scientist come GIGO "garbage in garbage out" per cui un algoritmo non può che riflettere la qualità dei dati su cui è costruito*" A. SIMONCINI, *Diritto costituzionale e decisioni algoritmiche, in (a cura di S. DORIGO) Il ragionamento giuridico nell'era dell'intelligenza artificiale*, Pisa, 2020, 60.

[35] In realtà il modello della ibridazione tra regole giuridiche e regole tecniche, mediante l'adozione di un "codice" vincolante i progettisti, è già utilizzato, proprio con riguardo al principio di spiegabilità degli algoritmi, dalla USACM (Associazione statunitense della meccanica computazionale), che, partendo dalla crescente evidenza che "alcuni algoritmi e analisi possono essere opachi, rendendo impossibile determinare quando i loro output possono essere distorti o errati prevede espressamente" e che "le decisioni prese dagli algoritmi predittivi possono essere opache a causa di molti fattori, tra cui tecnici (l'algoritmo potrebbe non prestarsi a una facile spiegazione), economici (il costo di fornire trasparenza può essere eccessivo, compreso il compromesso dei segreti commerciali) e sociali (rivelando l'input può violare le aspettative sulla privacy)" ha enucleato sette Principles for Algorithmic Transparency and Accountability, che vincolano la fase di progettazione e di sviluppo dei sistemi algoritmici, annoverando espressamente anche quello di spiegazione: "i sistemi e le istituzioni che utilizzano il processo decisionale algoritmico sono incoraggiati a fornire spiegazioni riguardanti sia le procedure seguite dall'algoritmo sia le decisioni specifiche che vengono prese. Ciò è particolarmente importante nei contesti di politica pubblicahttps://www.acm.org/binaries/content/assets/publicpolicy/2017_joint_statement_algorithms.pdf, visitato il 12 febbraio 2021.

[36] Nel *deep learning*, in particolare, non è possibile verificarne dall'esterno i meccanismi. In questi sistemi, il modello consiste nell'uso di"reti neurali artificiali" che, ad imitazione delle reti neurali umane, sono multistrato, nel senso che l'*output,* generato, sulla base di un *input*, ad un livello, viene inglobato e utilizzato come input nel livello successivo; e siccome ogni livello contiene migliaia di unità, i dati oggetto di elaborazione vengono continuamente elaborati nella loro interazione, con una complessità crescente che non è gestibile dall'intelligenza umana: in sostanza, di fatto l'algoritmo è capace di generare altri algoritmi. Per una descrizione comprensibile anche ai non esperti, cfr. A. LONGO-G.SCORZA, *Intelligenza artificiale. L'impatto sulle nostre vite, diritti e libertà,* cit., pag. 43-45.

l'opacità strutturale che caratterizza tali modelli è avvertita dagli stessi studiosi delle scienze informatiche, poiché uno dei campi di ricerca all'avanguardia dell'intelligenza artificiale, è quella della Explainable A.I. (XAI), volta a renderla comprensibile per l'essere umano, proprio per consentire che essa possa essere effettivamente utilizzata da chi poi è responsabile della decisione finale[37]; e ciò in base alla constatazione che, anche in termini di investimenti economici nello sviluppo, quale che sia la tecnologia adottata, se essa si presenta come una "scatola nera" finirà comunque con il trovare "resistenza" anche da parte di chi potrebbe adottarla con profitto.

Nel caso delle applicazioni di modelli di intelligenza artificiale nell'attività amministrativa, la resistenza però non è meramente operativa, ma normativa. La natura assoluta del principio di trasparenza, tale da prevalere anche su quelli di efficienza, economicità, semplificazione che governano il procedimento amministrativo ex art. 1 legge 241/90, costituisce, ad oggi, un vero e proprio filtro di ammissibilità del singolo modello, proprio alla luce dell'intelligenza artificiale antropocentrica, perseguita dalle istituzioni eurounitarie. In attesa che la ricerca in materia di Explainable A.I. (XAI) consegua i risultati sperati e consenta l'apertura delle "scatole nere", appare pertanto necessario che il principio della spiegabilità del modello algoritmo sia individuato a monte come uno standard di qualità, sia nell'ambito della valutazione comparativa delle diverse soluzioni disponibili imposta dall'art. 68 del d.lgs. 7 marzo 2005, 82 (CAD) [38]; sia nell'ipotesi in cui l'esigenza non possa essere soddisfatta ricorrendo a soluzioni già disponibili sul mercato e si ricorra al partenariato per l'innovazione ai sensi dell'art. 65 del D.Lgs. 18 aprile 2016, n. 50[39].

---

[37] Uno dei progetti in corso, avviati dal Laboratorio Nazionale di Artificial Intelligence and Intelligent Systems del CINI (Consorzio Interuniversitario Nazionale per l'Informatica), è "*HumanE AI Net*", finanziato dall'Unione Europoea: esso, attraverso la collaborazione i migliori centri di ricerca europei, le università e i poli industriali di eccellenza in materia di I.A., e la condivisione di una serie di competenze diverse (le scienze cognitive, le scienze sociali e la scienza della complessità) che ricalca il modello inclusivo già sperimentato anche agli albori dell'intelligenza artificiale, mira a rendere l'IA "comprensibile per l'uomo", lavorando sull'interazione uomo-macchina. Dal punto di vista della progettazione, l'idea è quella di inserire una serie di punti di controllo lungo tutto il processo di elaborazione, segmentare le operazioni in "moduli" alla fine dei quali il risultato è verificabile, e dove è la stessa macchina, cui è stato insegnato il linguaggio naturale, che spiega perché è arrivata a certe conclusioni: cfr. https://www.consorzio-cini.it/index.php/it/labaiis-home/labaiis-bandi/1628-progetto-humane-ai-net, visitato il 12 febbraio 2020.

[38] "*Le pubbliche amministrazioni acquisiscono programmi informatici o parti di essi nel rispetto dei princìpi di economicità e di efficienza, tutela degli investimenti, riuso e neutralità tecnologica, a seguito di una valutazione comparativa di tipo tecnico ed economico tra le seguenti soluzioni disponibili sul mercato: a) software sviluppato per conto della pubblica amministrazione; b) riutilizzo di software o parti di esso sviluppati per conto della pubblica amministrazione; c) software libero o a codice sorgente aperto; d) software fruibile in modalità cloud computing; e) software di tipo proprietario mediante ricorso a licenza d'uso; f) software combinazione delle precedenti soluzioni. 1-bis. A tal fine, le pubbliche amministrazioni prima di procedere all'acquisto, secondo le procedure di cui al codice di cui al decreto legislativo n. 50 del 2016, effettuano una valutazione comparativa delle diverse soluzioni disponibili sulla base dei seguenti criteri: a) costo complessivo del programma o soluzione quale costo di acquisto, di implementazione, di mantenimento e supporto; b) livello di utilizzo di formati di dati e di interfacce di tipo aperto nonché di standard in grado di assicurare l'interoperabilità e la cooperazione applicativa tra i diversi sistemi informatici della pubblica amministrazione; c) garanzie del fornitore in materia di livelli di sicurezza, conformità alla normativa in materia di protezione dei dati personali, livelli di servizio tenuto conto della tipologia di software acquisito*".

[39] "*1. Le amministrazioni aggiudicatrici e gli enti aggiudicatori possono ricorrere ai partenariati per l'innovazione nelle ipotesi in cui l'esigenza di sviluppare prodotti, servizi o lavori innovativi e di acquistare successivamente le forniture, i servizi o i lavori che ne risultano non può, in base a una motivata determinazione, essere soddisfatta ricorrendo a soluzioni già disponibili sul mercato, a condizione che le forniture, servizi*

o lavori che ne risultano, corrispondano ai livelli di prestazioni e ai costi massimi concordati tra le stazioni appaltanti e i partecipanti. 2. Nei documenti di gara le amministrazioni aggiudicatrici e gli enti aggiudicatori fissano i requisiti minimi che tutti gli offerenti devono soddisfare, in modo sufficientemente preciso da permettere agli operatori economici di individuare la natura e l'ambito della soluzione richiesta e decidere se partecipare alla procedura".