IA: lo human first e il processo amministrativo

Andrei Mihai Pop
Università degli Studi di Milano - Dottorando

Enrico Ferraris
PagoPA SpA

La rivoluzione dell'Intelligenza artificiale rappresenta uno spartiacque per la storia dell'Uomo. In particolare, uno dei settori che subirà influenze sarà quello della Pubblica amministrazione. Dopo aver preliminarmente definito il concetto di IA, si introduce il 'nuovo' diritto allo *human first*, che con la diffusione delle soluzioni basate su IA dovrà essere sempre più preminente. Infine, si analizza l'impatto che l'IA può avere sulla PA, in particolare nell'ambito del procedimento amministrativo, e come questa dovrà reagire per continuare ad operare nel rispetto dei diritti dei cittadini.

The Artificial intelligence revolution marks a watershed moment in human history. In particular, one of the sectors that will be significantly impacted is Public administration. After initially defining the concept of AI, the article introduces the 'new' right to what is known as 'human first', which, with the spread of AI-based solutions, will need to become increasingly prominent. Finally, the analysis focuses on the impact that AI may have on public administration, particularly in the context of administrative procedures, and how the latter must respond in order to continue operating in compliance with citizens' rights.

DOI: 10.53223/Sinappsi_2025-02-7

Citazione
Pop A.M., Ferraris E. (2025), IA: lo human first e il processo amministrativo, *Sinappsi*, XV, n.2, pp.90-100

Parole chiave
Diritti civili
Intelligenza artificiale
Pubblica amministrazione

Keywords
Civil rights
Artificial intelligence
Public administration

1. Una definizione europea di IA

Sul piano europeo, il principale contributo relativo alla definizione dell'IA è ascrivibile alle iniziative dei principali organi dell'Unione europea, la Commissione e il Parlamento. È la prima, infatti, che diede, inizialmente, l'impulso per la ricerca di una condivisa ed efficace definizione di IA. Il primo risultato di tali sforzi è contenuto in una Comunicazione intitolata *Artificial Intelligence for Europe* del 25 aprile 2018 nel quale le istituzioni europee manifestarono i propri obiettivi strategici in materia di IA, concentrandosi sulle opportunità e le sfide poste da questa tecnologia. In sintesi, questi possono essere suddivisi in due macro-tematiche: da un lato, il tentativo di divenire una potenza globale, incoraggiando gli enti pubblici e i soggetti privati a investire sull'IA, dall'altro la necessità di farlo nel quadro di un sistema legale ed etico che ponesse al centro del discorso una IA *trustworthy* e *human-centered*[1]. Tanto premesso, è opportuno sottolineare come questo primo approccio non rimase a lungo

Il presente lavoro è frutto di riflessioni condivise degli Autori, tuttavia il paragrafo 2 e 3 sono da ricondurre esclusivamente e unicamente ad Andrei Mihai Pop, mentre il paragrafo 1 ad Enrico Ferraris.

1 Si veda il paragrafo III di detto comunicato.

'operativo'. Nello specifico, infatti, può dirsi superato poco più di un anno dopo dai lavori dell'AI HLEG (un gruppo di esperti istituito dalla Commissione dell'UE nel 2018), il quale elaborò una nuova definizione di IA ad aprile del 2019[2]. Pur potendo considerare la fonte autorevole, questa definizione non ebbe effetti cogenti e vincolanti, non essendo considerabile come un atto legislativo. Invero, lo scopo degli esperti non era quello di regolare la materia, bensì di aprire un dibattito su tale nuova tecnologia. Il punto di svolta è rappresentato dalla proposta da parte della Commissione di un Regolamento, ossia l'AI Act, dell'aprile 2021[3], proposta che ha evidentemente tratto giovamento dal contributo del c.d. AI HLEG[4]. Il Regolamento, a seguito di una lunga gestazione incominciata fin dalla prima Comunicazione da parte della Commissione UE, venne approvato il 21/5/2024 e successivamente pubblicato il 12/7/2024 sulla Gazzetta ufficiale dell'UE (Regolamento UE, n. 1689/2024). In tale testo legislativo si elabora, a livello europeo, la prima definizione ufficiale e vincolante di IA. Infatti, l'articolo 3 comma 1 dell'AI Act statuisce che:

> *AI system means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments*[5].

Ai fini di una completa comprensione di tale definizione, è imprescindibile analizzare anche il *Considerando* n. 12 del Regolamento. In tale disposizione, il Legislatore dell'UE ha inteso chiarire alcuni aspetti che potrebbero rimanere 'incompresi' leggendo isolatamente l'articolo 3, comma 1, stante l'opacità connaturata nel linguaggio umano e la difficoltà definitoria della materia.

In primo luogo, in tale *Considerando*, si precisa che il Regolamento non si applica ai sistemi fondati unicamente su regole poste da persone fisiche che comportino l'esecuzione di operazioni in automatico. Questa è, a tutta evidenza, una formula che mira ad escludere dall'applicazione dell'AI Act alcune categorie di prodotti che non possono dirsi basati su Intelligenza artificiale. In tale prospettiva, il citato *Considerando* prevede che il Regolamento debba applicarsi esclusivamente in presenza di determinate caratteristiche tecniche. Invero, i sistemi di IA si fondano su almeno cinque caratteristiche che li differenziano dai tradizionali software: 1) capacità inferenziale; 2) approcci di apprendimento automatico; 3) approcci di logica; 4) autonomia; 5) adattabilità.

Ovviamente, la corretta e concreta declinazione di tali principi e regole dipenderà dalle interpretazioni che gli operatori del diritto ne daranno. Dunque, si può affermare che l'UE abbia deciso di favorire una omogenea prospettiva atlantista con la finalità di coordinarsi anche con gli orientamenti dei principali partner, specialmente gli Stati Uniti d'America. È evidente che riuscire a porsi tra le locomotive nella corsa allo sviluppo dell'IA dovrebbe poter garantire immensi vantaggi economici[6].

2 Si veda *Ethics Guidelines for Trustworthy Artificial Intelligence* (European Commission 2019).

3 La proposta è disponibile al seguente link: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 (consultato il 2/3/2025).

4 Purtuttavia, la definizione dell'AI Act non riproduce meramente le riflessioni dell'AI HLEG, bensì la reinterpreta, sviluppando e chiarendo alcuni aspetti. A titolo esemplificativo, si abbandona il riferimento tecnico al software o all'hardware preferendo la nozione di *machine-based system*.

5 L'Italia ha positivizzato – con la legge n. 132/2025 – una definizione, tuttavia è una mera riproposizione dell'indirizzo dell'UE non potendosi osservare alcuna differenza. Così anche Cassano (2024, 388).

6 L'approccio dell'AI Act mira ad avere gli stessi effetti che ebbe il GDPR. In tema di IA, si veda Finocchiaro (2024a, 113-125). Questo impatto dell'UE – prima col GDPR e ad oggi posto come obiettivo con l'AI Act – sugli altri attori internazionali è sintetizzabile con il c.d. *Brussels Effect*. Si veda Bradford (2020). Tuttavia, vi sono opinioni che ritengono più probabile un c.d. *Reverse 'effetto Bruxelles'*, ovverosia un forte rallentamento della crescita tecnologica nell'UE a causa dell'eccessiva regolamentazione. Inoltre, si sostiene che in realtà, almeno per quanto riguarda il 'mondo digitale' (contrapposto a quello 'fisico'), vi sarebbe non un 'effetto Bruxelles', bensì un 'effetto fotocopia'. In altre parole, l'Europa, in tale prospettiva, non riuscirebbe ad esportare la propria potenza normativa, ma assumerebbe solamente il ruolo di 'apripista' consentendo agli altri attori di trarre spunto, ferma però la possibilità di discostarsene. Si veda l'intervista del 26/8/2024 di Luciano Floridi disponibile al link: https://www.fondazioneleonardo.com/videos/effetto-bruxelles-normativa-europea-influenza-mondo-luciano-floridi (consultato il 4/3/2025); similmente Finocchiaro (2024b, 21-22). Per una prima analisi dei caratteri essenziali dell'AI Act – in particolare della governance – si rimanda anche a Pop (2024).

In definitiva, con questo AI Act, il vecchio continente mira a tracciare una 'terza via' che contemperi le esigenze economiche di sviluppo di una nuova tecnologia con le istanze etiche e giuridiche[7]. Una 'via europea' che si vorrebbe porre in alternativa rispetto all'impostazione legata al liberismo degli USA e a quella tipica dell'autoritarismo della Cina[8].

2. Lo human first come diritto fondamentale

Principiamo osservando quanto sia fondamentale porre al centro delle riflessioni, in tema dell'IA, l'essere umano. L'Uomo, con il suo ingegno, si è prodigato per migliorare le proprie condizioni di vita con l'uso-potenziamento della tecnica. Inizialmente, quest'ultima rimaneva strutturalmente uno strumento e mezzo asservito agli scopi del proprio creatore. Tuttavia, ad oggi, lo sviluppo della tecnica impone un sempre maggiore grado di autonomia delle invenzioni per incrementare il grado di soddisfazione della *societas*. Ciononostante, è doveroso tentare un contenimento di tale slittamento esistenziale attraverso lo strumento del diritto, il quale rimane una delle opere più longeve, potenti e 'umane'. La tecnica è presente nella vita dell'Uomo fin dai suoi albori: già l'uso di un bastone per procacciare cibo, o per offendere altri esseri, rappresentava la nascita di tale 'forza'. Purtuttavia, la *téchne* ha un gemello, nato in un periodo contiguo e, pressoché, in contemporanea. Ci si riferisce allo ius, il quale è già rinvenibile nelle prime regole di condotta di un dato gruppo di uomini rispetto ai propri membri (su un piano interno) e nei confronti di altri gruppi (su un piano esterno).

Queste due sono le 'forze' che indubbiamente sono presenti nella vita dell'Uomo fin dal principio e probabilmente lo accompagneranno per sempre. La tecnica come massima espansione delle ambizioni e potenze, alle volte incontrollate, dell'umanità; al contrario, lo ius come limite e guida verso un determinato scopo-fine che tenti di tutelare i diritti[9].

Negli ultimi decenni tale rapporto non è mutato, tuttavia la tecnica ha subìto una netta accelerazione e il diritto non è riuscito sempre ad adattarsi, anche a causa degli interessi economici che ne derivano e che impongono uno sviluppo sempre più rapido. Invero, sono addirittura sopraggiunte tesi che vorrebbero la sostituzione di uno ius 'umano' con un c.d. ius 'sintetico', in quanto oggettivo, imparziale e non più succube delle emozioni tipiche degli esseri umani. In tale ultimo scenario, la stessa esistenza del diritto, così come lo si conosce, rischierebbe di essere elisa[10]. Dunque, già da queste considerazioni iniziali, si può cogliere la necessità di respingere la tesi che vorrebbe un diritto assolutamente 'neutrale'[11] allo scopo di favorire il mercato, poiché una delle sue essenze è proprio quella di controbilanciare e di consentire la salvaguardia dei diritti degli individui. In tale prospettiva, quello che potrebbe essere ammesso è l'enucleazione di un insieme di norme coordinate per evitare 'fughe normative', ma non la sottomissione integrale dello ius alle esigenze mercantili. Questo parrebbe essere stato l'approccio dell'UE con l'AI Act emanato con l'intento di creare un mercato unico europeo anche nel campo dell'Intelligenza artificiale[12].

In tale contesto problematico si colloca l'emersione dell'IA, delle sue potenzialità e dei suoi rischi per i diritti fondamentali dell'Uomo. Questa tensione tra il dominio umano sulle proprie opere

7 Va però sottolineato che questa 'terza via' è una regolatoria e non produttivo-tecnologica. In altri termini, l'UE non ha la struttura e potenza di produzione necessaria per competere con gli Stati Uniti o con la Cina. Per tale ragione, l'UE sta tentando di spostare il 'campo di battaglia' sul piano normativo. Così Finocchiaro (2024a, 109).

8 Si veda Finocchiaro (2024a, 107-108). Questa impostazione mira a competere indubbiamente con l'America, ma specialmente con la Cina al fine evitare che il c.d. *Brussels Effect* venga sostituito da quello che è stato definitivo come il *Beijing effect*. Si rimanda per il *Beijing effect* a Erie e Streinz (2021).

9 Questa considerazione impone, altresì, che il diritto intervenga 'a monte' senza dover inseguire lo sviluppo costante e imprevedibile dell'Intelligenza artificiale. In altri termini, la portata e la forza del diritto deve inserirsi in un approccio by design all'IA che riesca a 'giocare d'anticipo' evitando, laddove possibile, di giungere al tradizionale approdo violazione-sanzione. Si veda Casonato (2019, 725).

10 Questa prospettiva pare non tenere sufficientemente in considerazione la circostanza per la quale, allo stato attuale, l'IA non sarebbe altro che un nostro riflesso, quindi con qualità, ma anche con difetti. Dunque, in tal guisa, la 'neutralità' dell'AI non sarebbe che un miraggio. Si veda Pietropaoli (2020, 110-113).

11 Sostiene tale tesi Abbott (2020, 134-143). Al contrario, Cal (2023, 65) vede nell'intervento del Legislatore europeo con l'AI Act un approccio "sostanzialmente neutro".

12 Il *Considerando* numero 1 dell'AI Act afferma che: "*The purpose of this Regulation is to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, the placing on the market, the putting into service and the use of artificial intelligence systems (AI systems) in the Union*".

e la ricerca continua di 'miglioramenti' della propria esistenza impone un contemperamento serio e stabile che solamente il diritto può svolgere.

In tale elaborato si accoglierà la tesi che impone la preferenza di un Uomo-dominus, rifiutando la deriva totale che impone di mutare la visione di un'umanità al centro del mondo, passando da una prospettiva tendenzialmente umano-centrica a una tecno-centrica. In tal guisa, la visione umano-centrica è stata cristallizzata anche dal Regolamento sull'Intelligenza artificiale che ha voluto imporre un'accelerazione su tali temi indicando una netta preferenza. Solamente in tal senso si può giustificare razionalmente l'enunciazione del diritto a una *AI Disclosure*, *Explainability* e *Transparency*. L'essere umano, dunque, deve mantenere un certo grado di controllo sull'IA e sul suo funzionamento al fine di tutelare i propri stessi diritti[13]. Questo criterio di gestione ha differenti gradi di 'potenza': la pervasività del controllo dell'Uomo rispetto alle operazioni compiute dall'IA aumenta all'aumentare dei diritti in gioco. Tali livelli di intervento sono conosciuti con le seguenti espressioni: *Human in command* (HIC), *Human in the loop* (HITL), *Human on the loop* (HOTL) e *Human out of the loop* (HOUTL)[14]. L'insieme di tali espressioni contribuisce alla formazione ed estensione del principio dello *human oversight*, ossia il controllo umano sul funzionamento dell'IA. Incominciando da quello più 'debole', ossia l'HOUTL, si può facilmente comprendere come non può trovare applicazione nel caso in cui siano in gioco i diritti delle persone: in questo scenario, infatti, vi è un'assenza totale dell'intervento umano, ovverosia il dominus assoluto è l'IA. Un simile livello di pervasività dell'IA appare, in via astratta, incompatibile con la maggior parte dei trattamenti che riguardino – direttamente o indirettamente – prerogative degli individui, salvo che i rischi siano nulli o minimi. Procedendo nell'analisi, e salendo nella 'scala dell'intervento umano', si giunge allo HOTL. In tale caso, l'IA ha l'iniziativa e la gestione del compito, mentre l'Uomo assume il ruolo di sorvegliare i suoi procedimenti potendo intervenire in una fase ex post rispetto all'output del sistema. Questo approccio può essere adottato laddove il sistema di IA non comporti pericoli rilevanti (rischi medio-bassi) per i diritti fondamentali. Invece, un grado di *human oversight* che consenta un più penetrante controllo-intervento dell'essere umano è quello dell'HITL che implica un controllo dell'essere umano dell'attività da svolgere: l'IA assiste, non decide. Questo scenario pare compatibile con la gestione della maggior parte dei casi d'uso che riguardino diritti fondamentali e che comportino un rischio medio-alto.

In ultimo luogo, si trova il c.d. HIC. In tale ipotesi, l'IA avrà un ristrettissimo ambito applicativo essendo l'intero procedimento controllato e guidato dall'Uomo. Contrariamente all'HOUTL, in tale caso è proprio l'Uomo ad essere il dominus. Tale rigore trova applicazione laddove le attività da svolgere possono coinvolgere e influenzare diritti fondamentali con un rischio alto-inaccettabile[15] – ad esempio, l'ambito della Giustizia. Dunque, il diritto allo *human oversight* deve essere declinato in base allo stato tecnologico e all'interesse oggetto di bilanciamento. Da un lato, più un sistema di IA è trasparente e spiegabile e maggiori sono le possibilità per seguire un approccio HOTL; dall'altro lato, assumendo una differente prospettiva, all'aumento di criticità del diritto oggetto di 'attenzione', vi è la necessità di utilizzare sistemi trasparenti e spiegabili. A prescindere dall''etichetta', tali espressioni possono essere ricondotte al diritto-dovere di applicare una sorveglianza umana sull'IA. Invero, questo criterio impone la regola secondo la quale in caso di utilizzo di tali sistemi, in settori giudicati come 'critici', occorrerà la supervisione dell'Uomo. In altri termini, l'Uomo e l'IA dovranno interagire e collaborare per giungere ad un output – c.d. *collaborative intelligence*[16]. Questo criterio-guida ha avuto già nel Regolamento europeo sulla protezione dei dati personali una sua iniziale, benché parziale, enunciazione all'articolo 22 comma 1, in materia di decisioni basate su trattamenti automatizzati. Ad una prima lettura, tale previsione risulta essere chiara e perentoria pur potendosi applicare solamente nei casi in cui l'IA utilizzi dati che possano essere qualificati come 'personali'. Inoltre, anche la prima parte del

13 In tali termini, pur riferendosi al c.d. elaboratore elettronico (potenzialmente dotato di IA), si veda Frosini (1983).

14 Propongono questa categorizzazione le *Ethics Guidelines for Trustworthy Artificial Intelligence* (European Commission 2019); per una prima interpretazione, si consenta di rimandare a Laviola (2020).

15 Abbiamo associato le varie categorie dell'*human oversight* con il rispettivo grado di rischio in base anche al *risk-based approach* dell'AI Act: HOUTL-nullo; HOTL-basso/medio; HITL-medio/alto; HIC-alto/inaccettabile.

16 In termini simili, Quintarelli *et al.* (2019, 191 e 199).

Considerando n. 71 del GDPR pare rinforzare tale prospettiva. Purtuttavia, tali enunciazioni subiscono importanti deroghe ai successivi commi, rendendo, de facto, il divieto di trattamenti automatizzati non la regola, bensì l'eccezione. Invero, il comma 2 dell'articolo 22 del GDPR prevede tre ipotesi in cui anche un trattamento automatizzato può essere considerato legittimo. In altri termini, il divieto non si applica laddove vi sia un contratto, una legge unionale o nazionale, ovvero nel caso in cui sia reso il consenso del soggetto destinatario di tale trattamento. In particolare, la deroga relativa all'accordo dell'interessato acuisce il rischio del c.d. *blind consent*, ossia la tendenza di prestare il proprio consenso, contrattuale o meno, ai fornitori dei servizi senza effettuare una sostanziale lettura dei termini e delle condizioni del trattamento (Marchetti e Casonato 2021). Tale tendenza diffusa deriva dal senso di urgenza di ottenere un rapido accesso ai servizi. Quindi, la previsione del GDPR che enuncia la necessità di un controllo umano è lacunosa per due ordini di ragioni: uno quantitativo e uno qualitativo. In primis, l'ambito applicativo riguarda solamente i dati personali e ciò esclude un gran numero di ipotesi in cui l'IA potrebbe operare; in secundis, le deroghe di cui al comma 2 dell'articolo 22 indeboliscono ulteriormente tale disposizione. Alla luce di quanto osservato, appare fondata la visione secondo la quale non si può ritenere esistente nel GDPR un diritto sufficientemente solido ed esteso al controllo umano nei trattamenti automatizzati da parte dell'IA.

Il Regolamento europeo sull'Intelligenza artificiale tenta di offrire una soluzione a tali aspetti problematici superando i limiti del GDPR. In particolare, una norma centrale è l'articolo 14 dell'AI Act – rubricato *Human oversight*. L'analisi di tale disposizione non può che cominciare dal primo comma. Infatti, qui si prevede che al fine di garantire la supervisione dell'Uomo sull'IA rientrante nella categoria 'ad alto rischio', quest'ultima dovrà essere ideata, progettata e realizzata in un modo compatibile con le capacità di comprensione e di gestione dell'essere umano. In altre parole, un sistema di IA ad alto rischio dovrà essere *human friendly by design*. La seconda parte di tale articolo fissa lo scopo di tale approccio, ovverosia quello di prevenire o ridurre i rischi per i diritti fondamentali delle persone oggetto del trattamento da parte dell'IA. Precisa la norma che tale valutazione, effettuata ex ante, deve tenere conto dell'uso 'naturale' di tale sistema, o al più di quello che si può ragionevolmente prevedere in base alle sue caratteristiche[17].

Tale precisazione è foriera di conseguenze, poiché richiama il tema della trasparenza e della spiegabilità: per riuscire ad anticipare un uso ragionevole del sistema occorre conoscere e comprendere il funzionamento dello stesso. In assenza di un sufficiente livello di tali caratteristiche, l'articolo 14 non si può essere rispettare. Già queste prime considerazioni sull'articolo rivelano l'intima connessione tra la *transparency*, la *explainability* e lo *human oversight*: essi sono tutti *conditio sine qua non*. Focalizzando l'attenzione sulla c.d. *explainability*, stante l'imponente sviluppo dell'IA, specialmente i sistemi più avanzati presentano una caratteristica unica e al contempo problematica, ossia l'essere una scatola nera (anche detta *black box*) (Pasquale 2015). In dottrina si discute del c.d. *black box problem*, il quale avrebbe origine da almeno due elementi: in primis, i Big Data che, a causa della loro stessa struttura, non consentono di ripercorrere il percorso decisionale della macchina; in secundis, l'apprendimento automatico – potenziato dalla tecnica del *deep learning* – che restituisce un sistema che prescindere da logiche deterministiche e quindi prevedibili. Dunque, la scarsa capacità esplicativa e l'imprevedibilità delle decisioni rendono l'IA portatrice di una nuova era di incertezza e di ignoto tecnologico (De Mari Casareto Dal Verme 2024, 3).

Questo significa che i passaggi logici che l'IA effettua, una volta che una determinata informazione è stata introiettata nel suo dataset, potrebbero risultare incomprensibili anche a un osservatore esperto. In altre parole, vi sarebbe una sensibile lacuna conoscitiva su quanto avviene nell'arco di tempo compreso tra l'input e l'output del sistema. Dunque, si può comprendere il contenuto dei dati di partenza e di quelli finali, tuttavia rimane oscuro il 'come' l'IA sia giunta a un certo risultato. In particolare, rimangono pressocché sconosciute le motivazioni che hanno indotto il sistema a modificare-manipolare i dati e le ragioni che hanno portato a un determinato esito (Pasceri 2023, 35).

17 Questo è il concetto di *reasonably foreseeable misuse*, il quale viene più volte valorizzato nell'AI Act all'articolo 3, n. 13.

Proseguendo, il comma 3 affronta il tema dell'intensità del controllo umano sul funzionamento del sistema di IA cercando di individuare un concreto contemperamento tra efficienza e diritti. La profondità della sorveglianza umana dipende dai diritti in discussione, dai rischi, dal livello di autonomia del sistema di IA e, infine, dal 'contesto' di riferimento e di funzionamento del sistema. Quindi, l'effettiva applicazione di tale regola deve tenere in considerazione tali quattro caratteristiche, in cui al mutare di una sola di esse si potrebbero verificare importanti scompensi imponendo la ricerca di nuovi equilibri. Il tema potrebbe essere concretizzato con un esempio. Si ipotizzi di voler applicare lo standard dello *human in command* alla luce di un diritto fondamentale in gioco dinanzi a un sistema di IA capace di porlo a un rischio elevato – rectius: con un grado di autonomia elevato – ma all'interno di un ambiente scientifico-sperimentale – c.d. *sandboxes*[18]. Quest'ultimo fattore risulta decisivo nel considerare non necessaria l'applicazione dell'HIC, poiché non vi sarebbe un 'reale' pericolo per il diritto.

Da ultimo, è opportuno osservare il comma 4. Quest'ultimo elenca le effettive misure che dovrebbero essere offerte all'utilizzatore finale per concretizzare il diritto al controllo umano sull'IA. Queste sono principalmente quelle legate alla possibilità di discostarsi dall'output prodotto dal sistema nella consapevolezza che lo stesso può essere errato, o comunque impreciso, alla luce dell'imperfezione del processo decisionale. Tale consapevolezza è cruciale e sarà uno dei punti più problematici nell'uso dei sistemi di IA ad alto rischio, poiché il rischio di accettazione acritica della decisione umana su quanto 'suggerito' dal sistema è concreto. In dottrina si è parlato di un effetto c.d. 'pecorone' (*anchoring effect*) (Garapon e Lassègue 2018, 239; Simoncini 2019, 81-83), ossia l'incapacità dell'Uomo di analizzare in maniera critica la decisione dell'IA in forza del mito dell'infallibilità della stessa. In altre parole, l'opzione individuata dal sistema diverrebbe la *default-option force* (Simoncini 2020, 39)[19], la quale godrebbe di un plusvalore di significato. Ovverosia, salvo palesi errori, per discostarsene servirebbe un intenso sforzo dell'essere umano che appare improbabile alla luce della forza persuasiva-psicologica del sistema. L'output dell'IA assume un simile valore poiché, dopotutto, riesce ad alleggerire il c.d. *burden of motivation*. Questa tematica sarà fondamentale perché solamente un vaglio 'umano' della decisione renderà quest'ultima 'accettabile-giustificabile' agli occhi dei consociati. Accanto alla possibilità di discostarsi dalla decisione dell'IA, l'AI Act aggiunge l'obbligo di inserire un comando d'emergenza che riesca a fermare l'attività del sistema di IA, ovvero di consentire l'intervento immediato dell'essere umano. Questo '*stop button*' ha sollevato polemiche da parte dell'industria che non lo giudica con favore.

Osservando l'impianto dell'articolo 14 dell'AI Act, si comprende l'ampiezza potenziale della sua applicazione, poiché i sistemi ad alto rischio occupano una percentuale importante dei sistemi di IA. Il rischio di un'interpretazione eccessivamente rigida potrebbe comportare alcuni problemi e il richiamo del comma 3 alle 'circostanze' di utilizzo dell'IA potrebbe non essere sufficiente per un 'equo' contemperamento tra le esigenze di tutela dei diritti e le necessità di sviluppo tecnologico e di competitività dell'UE. A tali dubbi daranno risposte gli interpreti, in particolare i giudici e le autorità (europee e nazionali) deputate al controllo circa l'attuazione del Regolamento. In conclusione, i nuovi diritti fondamentali che dovranno essere riconosciuti all'interno del sistema costituzionale italiano per consentire un governo 'umano-centrico' dell'IA sono rinvenibili sia nella legislazione europea sia nella legislazione nazionale. In tale capitolo ne abbiamo intravisti almeno tre, ossia *AI transparency*, *explainability* e *human first*.

Tali diritti fanno parte di una costellazione di principi intimamente legati tra di loro. Servirà una continua e proficua interazione tra gli stessi per consentire all'Europa di ergersi a 'faro mondiale', capace di concentrare i valori occidentali volti a un 'giusto' equilibrio tra l'ambizione della *téchne* e la tutela dell'Uomo. Tale contemperamento non può

18 L'articolo 3, n.55, dell'AI Act definisce la c.d. *sandbox* nei seguenti termini: "'*AI regulatory sandbox' means a controlled framework set up by a competent authority which offers providers or prospective providers of AI systems the possibility to develop, train, validate and test, where appropriate in real-world conditions, an innovative AI system, pursuant to a sandbox plan for a limited time under regulatory supervision*". In un simile contesto potrebbe essere derogata la normativa vigente al fine di consentire una determinata ricerca scientifica. A titolo esemplificativo, si veda il *Considerando* n.138 dell'AI Act.

19 Questa categoria deriva dai fondamenti della *nudging theory*. Si rimanda a Thaler e Sunstein (2021).

che essere frutto dell'intervento del diritto nella sua dimensione astratta, ma si evidenzia la necessità che le norme si trasformino da *law (right) in the books* in *law (right) in action* (Pound 1910). In caso contrario, non si potrà che ammettere l'abdicazione assoluta dello ius dal suo ruolo di co-protagonista – a fianco alla tecnica – e la sua riduzione a mera comparsa nella storia dell'umanità.

3. IA e Pubblica amministrazione

Alla luce dell'analisi svolta circa la genesi dei nuovi diritti in relazione all'avvento dell'IA, risulta opportuno cominciare a osservare gli effetti di una sua applicazione nel contesto della PA italiana. Questa scelta deriva dalla necessità di declinare il presente lavoro anche verso un ambito importante alla luce degli attori e dei diritti coinvolti.

Senza ombra di dubbio, la PA rappresenta un settore di grande interesse, ove l'applicazione dei sistemi di IA costituisce una delle tematiche più promettenti e, al contempo, complesse. La ratio di dedicare una particolare attenzione alla Pubblica amministrazione deriva dalla circostanza per cui, in tale ambito, il rapporto individuo-pubblico, e quindi la soggezione del cittadino a un potere statale, è particolarmente evidente.

In tempi recenti, proprio al procedimento amministrativo, dottrina e giurisprudenza hanno dedicato sempre maggiori riflessioni, nel tentativo di individuare un contemperamento tra la spinta dell'IA e la tutela dei diritti delle persone. Infatti, l'impiego di strumenti-sistemi di IA nella fase preparatoria di un provvedimento amministrativo e nella sua effettiva emanazione assume dei connotati critici che necessitano di attenta ponderazione e riflessione, poiché vi potrebbero essere conseguenze rilevanti per i cittadini. In altre parole, l'automazione (anche solo parziale) della decisione finale e della sua fase prodromico-valutativa erode l'idea di un potere pubblico 'antropocentrico' e può incrinare il necessario rapporto fiduciario che lega la PA al cittadino e che giustifica l'accettazione del privato di una decisione che lo riguardi. L'applicazione dei sistemi di IA in questo campo sta subendo un lento, ma costante incremento. Di conseguenza, sono sorte controversie circa la correttezza delle decisioni e delle modalità di impiego di tale tecnologia, le quali sono giunte sino al Consiglio di Stato. La sede giurisdizionale consente di assumere uno sguardo privilegiato, poiché facilita l'estrapolazione di quei principi-diritti che dovranno trovare concretizzazione nei futuri progetti.

Nel contesto italiano vi sono alcuni casi rilevanti in cui sono stati affrontati e analizzati gli algoritmi e il loro utilizzo all'interno della Pubblica amministrazione. Il primo è quello che riguarda la riforma della c.d. Buona scuola in forza della legge n. 107/2015.

In tale occasione, la sezione III-bis del TAR Lazio, con la sentenza n. 3769/2017, ha avuto l'occasione di affrontare la tematica dell'impiego degli algoritmi e delle procedure automatizzate all'interno della PA, avviando una riflessione in seno alla giurisprudenza amministrativa che ha subìto molteplici variazioni ed evoluzioni nel tempo. Questa pronuncia riconobbe che l'algoritmo, decidendo, de facto, le destinazioni degli insegnanti, si integra nel contesto di un procedimento amministrativo e, conseguentemente, dà origine al diritto all'accesso al codice sorgente e alle sottese logiche di funzionamento dello stesso ai sensi della legge n. 241/1990. Dunque si giunse alla conclusione che, alla luce dell'interesse pubblico in gioco, l'argomento della riservatezza dell'algoritmo, in quanto 'opera d'ingegno', non avrebbe potuto trovare accoglimento, non potendosi considerare come una causa legittima di esclusione ai sensi dell'articolo 24 della medesima legge. Inoltre, la circostanza che il Ministero avesse consegnato un documento che 'traduceva' in lingua italiana il linguaggio del programma, non era sufficiente ad escludere l'interesse a accedere direttamente a tale codice sorgente per poter verificare che la traduzione fosse effettivamente corretta. Solamente con tale modalità, infatti, è possibile controllare il funzionamento del programma e il suo rispetto della normativa vigente. Purtuttavia, e ben più rilevante ai fini del presente lavoro, il Tribunale – anche se solo *in obiter* – ha affermato che l'uso di un algoritmo decisionale all'interno di un procedimento amministrativo sarebbe stato valido e legittimo solamente a condizione che si fosse in un caso in cui fosse emerso l'esercizio di un c.d. potere amministrativo vincolato. Al contrario, laddove il contesto fosse stato quello di un c.d. potere amministrativo discrezionale, la conclusione sarebbe stata radicalmente opposta, ossia in senso negativo. Questa statuizione, benché *in obiter*, ha assunto crescente valore: invero alcune successive pronunce hanno ripreso tali argomentazioni trasformandole *in ratio decidendi,* costituendo così un terreno fertile per la nascita e consolidamento di un orientamento giurisprudenziale.

Le successive sentenze n. 9224/2018 e n. 9230/2018 del suddetto TAR Lazio hanno ribadito con vigore la necessaria natura servente dello strumento algoritmico, dovendo l'Uomo mantenere il controllo dei sistemi-procedure informatizzate e automatizzate. In questa prospettiva, un sistema di IA può al più divenire strumento ausiliario nelle mani del funzionario pubblico, dovendosi sempre garantire una qualche forma di 'dominio' sul procedimento amministrativo in ossequio agli articoli 7, 8, 10 della legge n. 241/1990 e 3, 24 e 97 della Costituzione. Dunque, in base a queste direttive interpretative, e anche nel contesto di una attività amministrativa vincolata, si può concludere che non vi può essere una sostituzione del contributo umano con un sistema automatizzato-artificiale, anche se si può arrivare a sostenere che più l'attività della PA è vincolata, minore deve essere il 'controllo umano'. Ciò senza, tuttavia, mai superare la *red line*, rappresentata dalla necessità di mantenere sempre l'Uomo come dominus.

Si comprende, a questo punto, l'atteggiamento di tale primo orientamento di apparente chiusura rispetto ai sistemi di IA automatizzati, poiché il cuore del procedimento amministrativo dovrebbe consistere nell'interazione e interlocuzione tra essere umani. In altri termini, benché la tecnologia possa 'servire' l'Uomo, si afferma il necessario 'volto umano' della PA.

Appare lapalissiano come tale impostazione teorica ricerchi di evitare un'automatizzazione della procedura amministrativa con conseguente spersonalizzazione della fase decisoria. Riprendendo la categorizzazione accennata nel paragrafo 2, parrebbe che tale orientamento non ammetta l'accesso all'HOTL e all'HOUTL, poiché richiede una prevalenza del contributo 'umano' e ciò già nel primo modello potrebbe essere mancante.

In tal guisa, si introduce un primordiale principio di non esclusività dell'attività 'artificiale' del sistema di IA, ovverosia il c.d. *human first* (benché limitato all'*Human in command* e all'*Human in the loop*), che è stato poi sancito definitivamente e con una portata 'maggiore' dall'articolo 14 dell'AI Act. Purtuttavia, benché vi siano degli interessanti spunti, questa prima riflessione dei giudici amministrativi poggia tralatiziamente sulla distinzione tra potere amministrativo vincolato e discrezionale. Quest'ultima però, come si dirà, non consente di impostare correttamente la riflessione, ovverosia di inglobare nel contemperamento dei principi-valori costituzionali anche quelli dell'efficienza, efficacia e buon funzionamento della Pubblica amministrazione. Questa conclusione, benché con l'intento di 'proteggere' l'Uomo, esclude aprioristicamente una riflessione proficua sull'uso dell'IA in un'ampia serie di ipotesi in cui, invece, ciò potrebbe avvenire sempre nel rispetto dei diritti delle persone.

La giurisprudenza amministrativa più recente ha ripreso i primi approdi dei giudici di primo grado chiarendo alcuni aspetti e ripensando alcune categorie. Un primo intervento dei giudici di Palazzo Spada[20], può essere rinvenuto nella sentenza della sezione VI del Consiglio di Stato, n. 2270/2019. In tale occasione, i giudici pur non superando l'asserita utilità della distinzione tra attività amministrativa vincolata e discrezionale, si sono concentrati su un altro aspetto, ossia l'auspicabile progressivo utilizzo dei sistemi di IA nella PA italiana.

L'impostazione iniziale nelle sentenze poc'anzi riportate aveva tentato di comprimere l'utilizzo dei sistemi automatizzati decisionali nel quadro delle attività amministrative vincolate. Dunque, pur ammettendo l'impossibilità di bloccare l'ingresso dell'IA nella PA, i giudici delle prime cure – ossia il primo grado innanzi al TAR – non avevano valorizzato appieno l'utilità della tecnologia, ammettendola solo in chiave residuale e in funzione esclusivamente ausiliare. Invece, il Consiglio di Stato, in questo primo intervento, si è espresso con favore all'introduzione di processi automatizzati all'interno del procedimento amministrativo, poiché ciò avrebbe anche aumentato l'efficacia e il buon funzionamento della PA in ossequio all'articolo 97 della Costituzione e all'articolo 1 della legge n. 241/1990. Inoltre, un simile *favor* nei confronti della tecnologia sarebbe stato conforme anche al Codice dell'Amministrazione digitale (D.Lgs. n. 82/2005) e agli influssi unionali. In altri termini, e specialmente laddove si fosse dinanzi a un gran numero di attività seriali-standardizzate in cui non fossero necessarie valutazioni discrezionali, l'implementazione di sistemi che consentano una gestione automatica avrebbe dovuto essere auspicabile e opportuna per aumentare la velocità dei procedimenti e per ridurre i casi di errori 'umani' dovuti a negligenza – oltre a una superiore garanzia circa l'imparzialità della decisione finale.

20 Con questa espressione si fa riferimento al Consiglio di Stato.

Alla luce di tali argomentazioni, si può osservare come i giudici abbiano inteso incentivare una declinazione della PA in chiave di e-government. Purtuttavia, va specificato che il Consiglio di Stato ha comunque ribadito la necessità che il funzionamento del sistema deve essere assolutamente conoscibile, sia al privato sia al giudice, in un eventuale giudizio. Questo tramite una traduzione chiara, completa ed effettiva del funzionamento del sistema. Dunque, la sentenza n. 2270/2019 del Consiglio di Stato, pur esprimendosi favorevolmente all'utilizzo di sistemi decisionali automatizzati in quanto coerenti con una PA che accoglie l'evoluzione tecnologica, ha mantenuto ferma la distinzione tra la categoria delle attività vincolate e quelle discrezionali. Cionondimeno, questo approdo del Consiglio di Stato non è quello definitivo. Infatti, è individuabile un'ulteriore evoluzione all'interno della sezione VI di Palazzo Spada con la sentenza n. 8472/2019. In tale occasione, i giudici hanno affermato che la suddetta distinzione non parrebbe più essere così attuale non potendosi, quindi, ritenere che l'utilizzo di tecniche automatizzate sia circoscritto alle sole attività amministrative vincolate. Invero, anche nelle attività connotate da discrezionalità potrebbe essere auspicabile l'introduzione di moduli organizzativi che utilizzino le potenzialità della c.d. Rivoluzione 4.0. Queste considerazioni nascono dall'idea secondo la quale l'elemento rilevante non è l'ambito di utilizzo degli algoritmi, bensì le cautele e i principi giuridici che trovano applicazione in chiave protettiva. Dunque, ciò a cui si deve guardare sarebbe il rispetto e la compatibilità di tali procedure automatizzate con le direttive generali del sistema giuridico. In ogni caso, benché in tale occasione si sia ammesso l'utilizzo degli algoritmi anche nel seno di un'attività a carattere discrezionale della PA, questo non implica una c.d. 'delega in bianco' all'amministrazione, dovendo essa sempre rispettare alcuni principi e regole fondamentali.

Purtuttavia, come ha segnalato autorevole dottrina, tale apertura alle c.d. attività amministrative discrezionali dovrebbe intendersi circoscritta tendenzialmente alle sole ipotesi di c.d. discrezionalità tecnica e non, invece, ai casi di c.d. discrezionalità 'pura'. In queste ultime ipotesi, le variabili e le circostanze da tenere in considerazione sono molteplici in base al caso concreto tanto che la stessa legge non regolamenta in toto tali circostanze lasciando il potere alla PA di valutare le singole situazioni. Per tali motivi, l'apertura giurisprudenziale dovrebbe essere limitata ai casi in cui troverebbe applicazione un basso grado di 'discrezionalità'.

Lo stesso Regolamento sull'IA con i suoi principi di trasparenza, spiegabilità e di *human oversight* parrebbe volgere in tale direzione imponendo un certo grado di controllo umano sull'attività dell'IA che aumenta in modo proporzionale in base alla rilevanza dei diritti in gioco.

Indubbiamente, come già anticipato, i giudici hanno ribadito l'opportunità di accogliere anche in seno alla PA la c.d. rivoluzione digitale, ma nel rispetto di alcuni principi fondamentali dell'ordinamento. In questa pronuncia, autorevole dottrina ha individuato l'enunciazione puntuale del principio di legalità algoritmica attraverso l'interpretazione del GDPR. Questa 'nuova' categoria impone che l'utilizzo da parte della PA di tali tecnologie avvenga nel rispetto dei generali principi costituzionali, europei e di quelli specifici del diritto amministrativo. In altre parole, sono le *conditio sine qua non* per delle decisioni automatizzate legittime da parte della PA.

Come si dirà di seguito, i principi individuati in tale occasione sono stati ripresi, rinforzati e arricchiti anche dall'AI Act, poiché il solo GDPR, come già evidenziato, non era sufficiente a garantire una piena tutela, a causa del suo ristretto ambito applicativo e delle numerose eccezioni che lo stesso prevede. I giudici, oltre che considerare gli articoli 13 comma 2 lettera f), 14 comma 2 lettera g), 15 comma 1 lettera h) e nell'articolo 22 comma 1, si soffermano anche sull'articolo 41 – rubricato "*Diritto ad una buona amministrazione*" –, comma 1 e 2, della Carta dei Diritti fondamentali dell'Unione europea (CDFUE).

In virtù di tale previsione, il principio di trasparenza algoritmica si declinerebbe nel senso che la PA, nel momento in cui si accinge a prendere una decisione che potrebbe incidere su un diritto di una persona, deve instaurare una interlocuzione (umana) prima di agire, con la conseguente necessità di mettere a disposizione le proprie informazioni e di motivare puntualmente le ragioni della propria determinazione.

Quindi, già da queste considerazioni si può intuire come la giurisprudenza, con la sua attività ermeneutica, abbia tentato e tenti di 'guidare' la transizione della PA nel quadro della rivoluzione tecnologica dell'IA. Tuttavia, con la sentenza n. 8472/2019 e con gli annessi richiami al GDPR, il Consiglio di Stato non si è limitato a un'analisi del principio di trasparenza, bensì ha tentato di andare oltre. In estrema sintesi, i giudici

si sono orientati al 'governo dell'algoritmo' inteso tout court, il quale deve essere, oltre che trasparente, anche sorvegliato dall'Uomo e non discriminatorio.

Circa la sorveglianza umana – ad oggi lo *human oversight* dell'articolo 14 dell'AI Act – i giudici di Palazzo Spada, riprendendo il modello dello *human in the loop* (con tutte le sue possibili varianti già viste), hanno affermato la necessità che il sistema interagisca proficuamente con l'essere umano non potendosi accettare procedimenti amministrativi in cui il dominus sia la macchina e non l'Uomo. Invece, per quanto concerne la non discriminazione algoritmica (i c.d. *algorithmic bias*), il Consiglio di Stato ha richiamato in particolare il *Considerando* 71 del GDPR. Una parte degli interpreti, in forza di tale riferimento, benché i giudici non avessero approfondito tale profilo, ha derivato anche il diritto al fatto che i dati utilizzati nella decisione siano corretti e di congrua 'qualità'. Purtuttavia, come si ha avuto modo di evidenziare, tale Regolamento – a causa del suo ambito applicativo, delle sue eccezioni e del valore giuridico dei *Considerando* nel quadro normativo europeo – non riusciva a garantire una piena e stabile tutela dinanzi ai sistemi di Intelligenza artificiale. Questo vulnus parrebbe essere stato colmato con l'AI Act, ma solo la giurisprudenza tramite la sua attività ermeneutica potrà dare risposte più precise su tali questioni.

In conclusione, dall'analisi di questa pronuncia del Consiglio di Stato, si può affermare che la giurisprudenza amministrativa abbia ricoperto un ruolo propulsivo non indifferente per un efficace contemperamento: da un lato, le esigenze di sviluppo e implementazione dell'IA in ossequio ai principi di efficacia, buon andamento ed economicità dell'azione della PA; dall'altro lato, quella relativa alla necessità che sia garantita la tutela dei diritti fondamentali delle persone sanciti da norme interne e unionali. Inoltre, tale orientamento appare ad oggi consolidato, poiché successivamente i giudici amministrativi sono tornati su tali argomenti specificando e ampliando alcuni punti e tematiche. Infatti, si può individuare almeno un altro significativo arresto giurisprudenziale, la sentenza del TAR Napoli, sez. III, n. 7003/2022. In tale occasione, i giudicanti hanno ribadito alcuni aspetti fondamentali riprendendo le due sentenze poc'anzi viste del Consiglio di Stato del 2019. In primis, il ricorso a una decisione c.d. algoritmica adottata in seno alla PA sarebbe auspicabile, in quanto idonea ad accrescere l'efficienza e il buon funzionamento dell'amministrazione pubblica. Inoltre, tale tecnologia non sarebbe applicabile solamente nel contesto di un'attività vincolata, bensì anche di una discrezionale. Ciò a condizione che sia garantito un certo grado di controllo umano – il TAR Napoli richiama la nozione di *Human in the loop* –, una congrua motivazione ed un certo grado di trasparenza. Su tale ultimo aspetto, è significativo come il Tribunale amministrativo regionale partenopeo espanda alcune considerazioni circa il principio di trasparenza in relazione all'uso di decisioni automatizzate della sentenza del Consiglio di Stato n. 8472/2019. Ovvero afferma che nella ricostruzione del 'significato' di un algoritmo e del suo funzionamento occorrono non solo competenze giuridiche, ma anche tecniche (ad esempio, in campo informatico e statistico). Di conseguenza, queste branche del sapere devono operare sinergicamente in modo che una 'regola tecnica' diventi una 'regola giuridica' comprensibile ai giudici e ai cittadini.

Conclusioni

Volendo compiere delle riflessioni finali e di sistema, appare essenziale evidenziare l'importanza dell'introduzione dell'IA in seno alla PA e del suo futuro sviluppo. Invero, tale nuova tecnologia dovrà essere inserita all'interno del contesto amministrativo al fine di favorire una c.d. buona amministrazione in ossequio al dettato costituzionale senza, però, trascurare il rispetto dei diritti fondamentali dei cittadini, anche mediante i citati principi dell'*human first*, della *transparency* e dell'*explainability*. Il volto 'antropocentrico' della PA rimane ad oggi un imperativo categorico incomprimibile, pertanto l'AI Act pare essere il miglior strumento a disposizione per tutelare tali aspirazioni.

Bibliografia

Abbott R. (2020), *The Reasonable Robot. Artificial Intelligence and the Law*, Cambridge, CUP

Bradford A. (2020), *The Brussels Effect: How the European Union Rules the World*, Oxford, OUP

Cal S. (2023), Il quadro normativo vigente in materia di IA nella pubblica amministrazione, in Belisario E., Cassano G. (a cura di), *Intelligenza artificiale per la pubblica amministrazione. Principi e regole del procedimento amministrativo algoritmico*, Pisa, Pacini Editore, cap.3

Casonato C. (2019), Costituzione e intelligenza artificiale: un'agenda per il prossimo futuro, *Rivista di Biodiritto*, n. speciale 2, pp.711-725

Cassano G. (2024), Note minime sul D.D.L. in materia di Intelligenza Artificiale, *Diritto di Internet*, 28, n.3, pp.579-590

De Mari Casareto dal Verme T. (2024), *Intelligenza artificiale e responsabilità civile. Uno studio sui criteri di imputazione*, Collana della Facoltà di Giurisprudenza dell'Università di Trento n.48, Napoli, Editoriale Scientifica

Erie M.S., Streinz T. (2021), The Beijing Effect: China's Digital Silk Road as Transnational Data Governance, *Journal of International Law and Politics*, 54, n.1, pp.1-91

European Commission (2019), *Ethics guidelines for trustworthy AI*, Brussell, European Commission

Finocchiaro G. (2024a), *Intelligenza artificiale. Quali regole?*, Bologna, il Mulino

Finocchiaro G. (2024b), *Diritto dell'intelligenza artificiale*, Bologna, Zanichelli

Frosini V. (1983), L'informatica e la pubblica amministrazione, *Rivista trimestrale di diritto pubblico*, 33, n.2, pp.483-494

Garapon A., Lassègue J. (2018), *Justice digitale. Révolution graphique et rupture anthropologique*, Paris, Presses Universitaires de France

Laviola F. (2020), Algoritmico, troppo algoritmico: decisioni amministrative automatizzate, protezione dei dati personali e tutela delle libertà dei cittadini alla luce della più recente giurisprudenza amministrativa, *Rivista di Biodiritto*, n.3, pp.389-440

Marchetti B., Casonato C. (2021), Prime osservazioni sulla proposta di Regolamento dell'Unione europea in materia di Intelligenza artificiale, *Rivista di Biodiritto*, n.3, pp.415-437

Pasceri G. (2023), Le scienze argomentative tra stereotipi e veri pregiudizi: la 'black box', *Ciberspazio e Diritto*, 24, n.73, pp.23-44

Pasquale F. (2015), *The Black-Box Society: The Secret Algorithms That Control Money and Information*, Cambridge (MA), Harvard University Press

Pietropaoli S. (2020), Fine del diritto? L'intelligenza artificiale e il futuro del giurista, in Dorigo S. (a cura di), *Il ragionamento giuridico nell'era dell'Intelligenza Artificiale*, Pisa, Pacini Giuridica, pp.101-120

Pop A.M. (2024), Il risk-based approach, la governance e le sanzioni nell'AI Act, *Ciberspazio e Diritto*, 25, n.3, pp.423-436

Pound R. (1910), Law in Books and Law in Action, *American Law Review*, n.34, pp.12-36

Quintarelli S., Corea F., Fossa F., Loreggia A., Sapienza S. (2019), AI: profili etici. Una prospettiva etica sull'Intelligenza Artificiale: principi diritti e raccomandazioni, *Rivista di Biodiritto*, 3, pp.183-204

Simoncini A. (2020), Amministrazione digitale algoritmica. Il quadro costituzionale, in Perin R.C., Galetta D.-U. (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, pp.33-52

Simoncini A. (2019), L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà, *Rivista di Biodiritto*, n.1, pp.63-89

Thaler R.H., Sunstein C.R. (2021), *Nudge. Improving Decisions About Health, Wealth, and Happiness*, London, Penguin Publishing Group

Andrei Mihai Pop
andrei.pop@unimi.it
È dottorando di ricerca presso il Dipartimento di Scienze giuridiche Cesare Beccaria dell'Università degli Studi di Milano. Fra le pubblicazioni recenti si segnala: Il risk-based approach, la governance e le sanzioni nell'AI Act, *Ciberspazio e Diritto*, n.78, 2024.

Enrico Ferraris
enrico.ferraris@pagopa.it
È un avvocato specializzato in diritto delle nuove tecnologie con un focus su protezione dei dati, privacy e Intelligenza artificiale. Attualmente ricopre il ruolo di Responsabile Product privacy & ethics presso PagoPA SpA, dove coordina l'implementazione dei principi di *privacy by design* nei prodotti ed *ethics by design* delle soluzioni basate su IA. È attualmente membro del comitato direttivo del Dottorato di ricerca in Giurisprudenza presso l'Università di Milano ed esperto esterno del Comitato etico per le tecnologie emergenti del Comune di Torino.