

1/2026 anteprima

# Rivista di diritto dei media

ISSN 2532-9146

# Tra predizione e protezione dei dati: la pubblica amministrazione algoritmica al vaglio dei rischi e delle garanzie

Riccardo Merli

## Sommario

1. Premessa. - 2. Dalla digitalizzazione amministrativa alla *governance* algoritmica. - 3. Dati personali e nuove architetture informatiche: un'intersezione problematica. - 3.1. Opacità algoritmica e motivazione. - 3.2. Il consenso come base fragile. - 3.3. Responsabilità amministrativa e *accountability* progettuale. - 3.4. Profilazione, *bias* e principio di eguaglianza. - 4. L'art. 22 GDPR: decisioni automatizzate e diritto alla spiegazione. - 5. Conclusioni.

## 1. Premessa

Viviamo in una fase storica segnata da una rivoluzione informatica sistemica, in cui la Pubblica Amministrazione (PA) ha progressivamente abbandonato la dimensione meramente documentale per rivestire un ruolo strategico nell'adozione e nello sviluppo delle nuove tecnologie<sup>1</sup>. A ben vedere, infatti, la digitalizzazione, inizialmente limitata alla dematerializzazione dei procedimenti e alla costruzione della cittadinanza digitale, si è presto tradotta in forme più sofisticate di automazione e predizione, capaci di incidere in modo diretto sul rapporto tra amministrazione e cittadini<sup>2</sup>. In questo scenario complesso e in continua trasformazione, i giuristi sono chiamati a confrontarsi con questioni in larga parte inedite ma ormai inderogabili. Ciò che emerge, infatti, non è un mero problema di modernizzazione tecnica della macchina amministrativa, bensì una sfida che investe principi costituzionali e categorie giuridiche consolidate che sono spinte a verificare la loro adattabilità alla tecnologia.

---

<sup>1</sup> I. D'Elia Ciampi, *L'informatica e le banche dati*, in S. Cassese (a cura di), *Trattato di diritto amministrativo, Parte Speciale*, Milano, 2003, IV, 1632.

<sup>2</sup> Per una ricostruzione sulle fasi della digitalizzazione e sulle principali disposizioni in materia *ex multis* vd. P. Rubecchini, *Tecnologia blockchain e fiducia amministrativa*, Napoli, 2023; G. Pesce, *Digital first*, Napoli, 2018.

In altri termini, la progressiva introduzione di sistemi algoritmici<sup>3</sup> non si limita a ridefinire le modalità operative degli uffici ma, sostituendo la logica tradizionale del controllo *ex post* con quella del condizionamento *ex ante*, ha inciso sul terreno della legalità, della trasparenza, della responsabilità e, in special modo, dell'eguaglianza sostanziale nell'accesso ai diritti e alle prestazioni pubbliche<sup>4</sup>.

Da questa prospettiva, l'efficienza algoritmica diviene una nuova forma di legittimazione dell'azione pubblica, ma – come si mostrerà nel prosieguo – essa non è neutra<sup>5</sup> né priva di tensioni con i principi costituzionali<sup>6</sup>.

A partire da tali riflessioni, il percorso di ricerca che qui si inaugura si snoda lungo tre direttrici complementari che intendono ricostruire un quadro dogmatico unitario del rapporto tra predizione e protezione dei dati.

La prima direttrice riguarda la ricostruzione del mutamento concettuale che ha condotto dall'amministrazione digitale all'amministrazione predittiva, segnando il passaggio da un modello informativo a uno propriamente cognitivo, fondato sull'uso sistematico dei dati a fini decisionali.

La seconda, ragionando sul principio costituzionale di eguaglianza, analizza come *bias*, opacità e processi di profilazione possano tradursi in nuove forme di diseguaglianza sostanziale e, quindi, come la responsabilità amministrativa possa divenire una condizione di giustizia digitale, *id est* quale nuovo assetto di garanzie dell'azione amministrativa algoritmica.

La terza, infine, concerne l'esame delle garanzie previste dal quadro normativo europeo, con particolare attenzione all'art. 22 del Regolamento (UE) 2016/679, anche conosciuto come *General Data Protection Regulation* (GDPR)<sup>7</sup>, e la loro interazione con gli obblighi di supervisione umana introdotti dal Regolamento (UE) 2024/1689, altresì noto come Regolamento sull'intelligenza artificiale (d'ora in poi, "AI Act"), al fine di verificare limiti e potenzialità nella disciplina delle decisioni automatizzate.

Ciò chiarito, occorre precisare che l'indagine è basata su approccio giuridico-dogmatico, che non ambisce a fornire risposte conclusive in grado di porre la parola "fine". Essa si propone, piuttosto, di offrire una ricostruzione critica volta ad alimentare il dibattito e a verificare se, e in quale misura, sia possibile coniugare innovazione tecnologica e tutela dei diritti fondamentali.

---

<sup>3</sup> Per un inquadramento giuridico degli algoritmi nell'attività amministrativa si veda A. Sola, *Inquadramento giuridico degli algoritmi nell'attività amministrativa*, in *Federalismi.it*, 16, 2020, 332 ss.

<sup>4</sup> Non a caso, la dottrina ha da tempo messo in luce ritardi e disomogeneità, rilevando come l'assenza di standard uniformi dei servizi digitali rischi di compromettere l'eguaglianza sostanziale nell'accesso alle prestazioni pubbliche. Sul punto si veda M. Pietrangelo, *Sui "diritti di cittadinanza digitale"*. Note a margine di un opaco percorso normativo, in *Federalismi.it*, 8, 2022, 130 ss.

<sup>5</sup> Tra i molti vd. A. C. Amato Mangiameli, *Algoritmi e big data. Dalla carta sulla robotica*, in *Rivista di filosofia del diritto*, 1, 2019, 107 ss.

<sup>6</sup> Tant'è vero che in dottrina, in materia di algoritmi, si è proposta una dottrina della "precauzione costituzionale". Così A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal*, 1, 2019, spec. 86 e 87.

<sup>7</sup> In relazione al quadro delineato dal GDPR in materia si vd. R. C. Perin-D. U. Galetta, *Il Diritto dell'Amministrazione Pubblica digitale*, Torino, II, 2025, spec. 60 ss.

---

## 2. Dalla digitalizzazione amministrativa alla governance algoritmica

Appare doveroso chiarire che l'evoluzione digitale della PA non può essere ridotta a una sequenza lineare di fasi settoriali. Piuttosto, essa si è venuta a configurare come un processo di trasformazione complessiva che, tanto sul piano dogmatico quanto su quello descrittivo, segna il passaggio da una digitalizzazione debole a una digitalizzazione forte, capace di investire in profondità ogni processo e dimensione dell'agire pubblico<sup>8</sup>.

La prima si concentra sull'estetica e sull'organizzazione dell'attività amministrativa, ossia su quell'insieme di interventi che attengono all'esteriorità del procedimento. Ne sono un esempio la dematerializzazione dei documenti e, con essa, l'interazione online con cittadini e imprese tramite moduli elettronici, firme digitali, portali, protocolli informatici e servizi di posta elettronica certificata.

Messi in chiaro questi profili, appare evidente che il passaggio da una PA cartacea a una "paperless"<sup>9</sup> rappresenta certamente un traguardo significativo, che però rimane pur sempre limitato ad una mera trasposizione in formato digitale di procedure e archivi analogici, senza incidere sul nucleo cognitivo dell'agire pubblico. Al contrario, la digitalizzazione forte rappresenta un'autentica ristrutturazione epistemica che ha segnato l'arrivo di un "nuovo paradigma amministrativo"<sup>10</sup>.

In questa fase, la tecnologia non si limita più a svolgere un ruolo meramente ausiliario, ma diventa parte integrante e determinante dell'azione amministrativa, fino a incidere sulla sua stessa fisionomia. Ed è all'interno di questa nuova configurazione epistemica che trovano spazio strumenti di analisi avanzata e tecniche di *machine learning*; vale a dire, sistemi che non si limitano a riportare asetticamente la realtà, ma la producono e la orientano, modellando comportamenti e decisioni amministrative secondo logiche probabilistiche.

Sotto questa prospettiva, infatti, parte della dottrina ha parlato di una vera e propria "secessione amministrativa"<sup>11</sup>, espressione che ben restituisce l'idea di una progressiva retrocessione del ruolo umano all'interno del procedimento; *scilicet*, il cittadino non è più destinatario del provvedimento ma oggetto di una classificazione probabilistica. Un processo che, come avremo modo di dire, ha reso presto evidente l'esigenza di nuove garanzie. Per comprendere appieno la portata di tale fenomeno, è necessario chiarire

---

<sup>8</sup> G. Duni, *Amministrazione digitale*, in *Enciclopedia del diritto*, I, 2007, 17, invece, parla di passaggio da un «momento statico» a un «momento dinamico».

<sup>9</sup> In altri termini, la digitalizzazione implica il venir meno del legame tradizionale con il supporto cartaceo, che resta confinato a un ruolo meramente accessorio rispetto allo svolgimento dell'attività amministrativa. Cfr. G. Duni, *La amministrazione digitale. Il diritto amministrativo nell'evoluzione telematica*, Milano, 2008, 14.

<sup>10</sup> G. Carullo, *La nozione di servizi digitali: un nuovo paradigma per la pubblica amministrazione autocertificazione e diritto dell'Unione europea*, in *Istituzioni del federalismo: rivista di studi giuridici e politici*, 2, 2023, 335 ss.

<sup>11</sup> Sul punto cfr. D. Diaco, *Amministrazione umana vs Amministrazione algoritmica: prolegomeni su un modello procedimentale a partecipazione successiva*, in *Judicium. Il processo civile in Italia e in Europa*, 2023.

che cosa significhi, in termini tecnici, “predire”: elaborare inferenze basate su tecnologie statistiche e algoritmiche, in particolare mediante l’ausilio dell’intelligenza artificiale (IA), capaci di anticipare eventi futuri osservabili direttamente a partire dall’analisi di dati passati o presenti.

A ben vedere, pertanto, contrariamente ai metodi statistici tradizionali, i sistemi cognitivi non si affidano a una sequenza rigida di calcoli deterministici<sup>12</sup>, né operano secondo regole predefinite che ne garantiscono la tracciabilità, ma utilizzano logiche più flessibili, talvolta *fuzzy*<sup>13</sup>, più adatte a gestire l’incertezza e, proprio per questo, potenzialmente più accurate e stringenti. Esse si fondano, invero, su modelli di apprendimento automatico (appunto, *machine learning*) in grado di evolvere «in maniera proporzionale rispetto al numero di dati che vengono elaborati»<sup>14</sup> e che, processando tali grandi volumi informativi (*big data*<sup>15</sup>), sono in grado di individuare *pattern*, correlazioni e segnali all’interno del “rumore informativo”, restituendo, così, un risultato probabilistico deduttivo scarsamente bisognoso di un mero assenso umano.

In buona sostanza, l’*ubi consistam* del sistema automatizzato, a differenza di quello tradizionale, limitato a riportare relazioni osservate nei dati storici, risiede nella sua capacità di applicare inferenze apprese in fase di addestramento a situazioni nuove e non perfettamente coincidenti con quelle già conosciute, dando luogo a un ragionamento probabilistico che, pur non assicurando esiti certi, consente di orientare l’azione amministrativa in scenari complessi e mutevoli, adeguando la risposta alle nuove variabili. Per comprendere appieno il funzionamento di tali strumenti, è forse utile osservare in quali procedimenti essi trovano oggi applicazione concreta nel dominio amministrativo; per esempio, essi vengono impiegati per valutare il rischio di evasione dei contribuenti, per stimare la probabilità di recidiva di un imputato, per il riconoscimento dei soggetti più esposti all’emarginazione sociale e per individuare i destinatari più appropriati di benefici o interventi pubblici.

Questa evoluzione mostra, infatti, come, nel tempo, il passaggio da un modello burocratico e reattivo a un’amministrazione che si configura come organismo cognitivo potenziato, ha di contro affermato un’autentica *governance* del dato<sup>16</sup>. Il dato stesso diviene il perno della vita amministrativa,

---

<sup>12</sup> Francaviglia distingue tra “informatizzazione della P.A.”, “elaborazione algoritmica – e dunque automatizzata – del provvedimento amministrativo” e “attività provvedimentoale *data-driven*”. Nello specifico, gli algoritmi predittivi rientrano principalmente nelle ultime due categorie. Così M. Francaglia, *L’intelligenza artificiale nell’attività amministrativa: principi e garanzie costituzionali nel passaggio dalla regola agendi alla regola algoritmica*, in *Federalismi.it*, 17, 2024, 114 ss.

<sup>13</sup> Per approfondire cfr. M. Veronesi-A. Visioli, *Logica fuzzy. Fondamenti teorici e applicazioni pratiche*, Milano, 2003.

<sup>14</sup> G. Alpa, *Diritto e intelligenza artificiale*, Pisa, 2020, 282.

<sup>15</sup> *Ex multis* cfr. C. Perin (a cura di), *L’amministrazione pubblica con i big data: da Torino un dibattito sull’intelligenza artificiale*, in *Quaderni del Dipartimento di Giurisprudenza dell’Università di Torino*, Torino, 2021.

<sup>16</sup> Per un approfondimento sulla *data governance* e sulla sovranità digitale nella PA, cfr. V. Pagnanelli, *Conservazione dei dati e sovranità digitale. Una rilettura della (big) data governance pubblica alla luce delle nuove sfide globali*, in *Rivista Italiana dell’Informatica e del Diritto*, 1, 2021, 11 ss.

trasformando le pubbliche amministrazioni in vere e proprie *data-driven organizations*<sup>17</sup>, nelle quali gli strumenti per la raccolta, l'elaborazione e la gestione delle informazioni non ricoprono più un ruolo di mero supporto esterno all'*iter* amministrativo, ma tendono a diventare ontologicamente il processo stesso.

Lumeggiate la funzione e il fine di questi strumenti, nel prossimo paragrafo occorrerà constatare come la loro applicazione sollevi rilevanti profili problematici di ordine giuridico e costituzionale, imponendo di soffermarsi sui rischi connessi e sulle diverse forme di regolazione che il legislatore è chiamato ad approntare. La questione, infatti, non è se usare o meno algoritmi, ma a quali condizioni il loro impiego sia compatibile con i principi dell'azione amministrativa e con le tutele del GDPR.

### **3. Dati personali e nuove architetture informatiche: un'intersezione problematica**

Nei paragrafi precedenti abbiamo avuto modo di constatare come all'inizio del XXI secolo si è affermato un fenomeno noto come *datafication*<sup>18</sup>, alcuni hanno parlato persino di “datacrazia”<sup>19</sup>, ossia la tendenza a trasformare ogni aspetto della realtà in dati quantificabili, capaci di guidare, influenzare o addirittura determinare le decisioni politiche, economiche, amministrative o sociali, spesso attraverso strumenti automatizzati come algoritmi e intelligenza artificiale.

In virtù di ciò, insomma, i dati sono oramai considerati una risorsa strategica a livello globale, tanto da essere spesso descritti come il “nuovo petrolio”<sup>20</sup>. Tuttavia, l'analogia è solo parziale. A differenza delle risorse

---

<sup>17</sup> Cfr. A. Pentland, *Social Physics: how good ideas spread-the lessons from a new science*, New York, 2014; D. Kiron, *Lessons from becoming a data-driven organization*, in *MIT sloan management review*, 58-2, 2017; P. Korherr-D. K. Kanbach-S. Kraus-P. Mikalef, *From intuitive to data-driven decision-making in digital transformation: A framework of prevalent managerial archetypes*, in *Digital Business*, 2, 2022.

<sup>18</sup> Secondo S. Calzolaio (*Introduzione. Ubi data, ibi imperium: il diritto pubblico alla prova della localizzazione dei dati*, in *Rivista italiana di informatica e diritto*, 1, 2021, spec. 5) la *datafication* si fonda su tre elementi: la crescita esponenziale della produzione di dati, la capacità delle macchine di analizzarli ed estrarne conoscenza e, infine, l'utilizzo di tali informazioni nei processi decisionali tramite *algorithmic decision making*. Per un ulteriore approfondimento del tema, *ex multis*, si veda K. Cukier-V. Mayer-Schöenberger, *Big data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, Milano, 2013; C. Sarra, *La dataficazione della persona nella costruzione del Metaverso*, in *Journal of Ethics and Legal Technologies*, 6, 2024, 3 ss.

<sup>19</sup> S. Uggeri, *Il potere dei Big Data e la datacrazia della rete*, in *Confronti: mensile di fede, politica, vita quotidiana*, 1, 2018, 29 ss.; D. Gambetta (a cura di), *Datacrazia: Politica, cultura algoritmica e conflitti al tempo dei big data*, Roma, 2018; D. Talia, *Il rischio della Datacrazia nelle Smart Cities*, in G. F. Ferrari (a cura di), *Innovazione e sostenibilità per il futuro delle smart cities*, Milano, 2023, 793 ss.

<sup>20</sup> B. Ando, *Dati personali tra Costituzione e mercato: una prospettiva di diritto comparato*, in M. Mazzuca-M. Mauro (a cura di), *Sulle interrelazioni tra settori del sistema giuridico. Esperienze interdisciplinari*, Napoli, 2024, 61 ss.; V. Montozzi, *Le nuove sfide della regolazione dei dati: analisi di un modello a partire dall'intersezione tra protezione dei dati personali e concorrenza*, in *Rivista italiana di informatica e diritto*, 1, 2025, 501 ss.

naturali esauribili, i dati non si consumano, ma possono essere copiati, condivisi, riutilizzati indefinitamente e, persino, venduti (all'uopo, si tornerà più avanti sul caso *Weople*), mantenendo intatto il loro valore d'uso. Ciò che li rende realmente produttivi è, piuttosto, la loro organizzazione e trattamento; *in nuce*, come il greggio deve essere raffinato per diventare carburante, così i dati necessitano di essere raccolti, selezionati, correlati e analizzati.

In questa prospettiva, l'intelligenza artificiale, come già spiegato, svolge un ruolo decisivo, consentendo l'elaborazione di inferenze e previsioni che superano i limiti cognitivi dell'uomo. E, tuttavia, ridurre l'uso degli algoritmi a una dimensione salvifica significherebbe trascurarne le criticità e i rischi strutturali.

È indubbio, infatti, che la credibilità complessiva della ricostruzione qui delineata resta condizionata dal riconoscimento che l'algoritmo, allo stato attuale, non è in alcun modo qualificabile come un operatore o, peggio, come un giudice infallibile, bensì come un mero moltiplicatore di capacità classificatoria e predittive già presenti nei dati. Esso, invero, non sostituisce la razionalità umana, ma la moltiplica, offrendosi come strumento di elaborazione capace di ridurre i tempi procedurali, favorire l'uniformità di trattamento, rendere misurabili le *performance* e tracciabili gli esiti. Eppure, è proprio in questo processo che si coglie la portata innovativa dei modelli computazionali, la cui attrattiva rimane innegabile.

Si badi. Il disquisito ecosistema digitale, pur tra le sue ambivalenze – che, per certi versi, richiama il chiaroscuro caravaggesco, ove la luce e l'ombra convivono in una tensione drammatica volta a guidare lo sguardo sui nuclei essenziali della scena –, non deve indurre in confusione. È doveroso, infatti, tener presente che la complessità del fenomeno, nella sua effettiva portata, impone al presente saggio un'alternanza di considerazioni di segno positivo e negativo, in un equilibrio instabile in cui non è agevole determinare l'inclinazione del piatto della bilancia.

In buona sostanza, se sul piano generale l'innovazione apre scenari di straordinaria portata, sul piano più concreto emergono con forza le problematiche connesse al rapporto tra dati personali e nuove architetture informatiche<sup>21</sup>.

È noto, infatti, come anche la cronaca più recente abbia mostrato gli enormi rischi che si accompagnano alla centralità dei dati. Per fare qualche esempio, nel 2024 la Direzione Distrettuale Antimafia di Milano ha portato alla luce una rete di spionaggio informatico che, mediante accessi abusivi a banche dati pubbliche e private, metteva in commercio informazioni riservate di centinaia di migliaia di cittadini e imprese<sup>22</sup>. Pochi mesi più tardi, un'indagine parallela a Bari ha rivelato analoghe pratiche di intrusione

---

<sup>21</sup> Sui profili di incompatibilità delle tecniche di polizia predittiva con la tutela di *privacy* e dati personali vd. A. Bonfanti, *Big data e polizia predittiva: riflessioni in tema di protezione del diritto alla privacy e dei dati personali*, in *Media Laws Rivista di diritto dei media*, 3, 2018, 1 ss.

<sup>22</sup> L'inchiesta milanese ha coinvolto la società investigativa *Equalize*, guidata dall'ex capo della Polizia Carmine Gallo, accusata di aver creato un vero e proprio mercato nero dei dati, attingendo abusivamente da banche dati come SDI, Serpico, INPS, ANPR e SIVA. Tra i clienti figuravano grandi imprese italiane ed estere, che pagavano somme tra i 1.000 e i 15.000 euro per l'ottenimento di informazioni riservate.

sistematica ai danni dei correntisti da parte di un ex dipendente bancario, tra i quali figuravano anche esponenti delle massime istituzioni<sup>23</sup>. Questi episodi mettono in evidenza non solo l'esistenza di un vero e proprio mercato nero dei dati, ma soprattutto le carenze strutturali dei sistemi di protezione, con i rischi di furto, di alterazione e/o di distruzione dei dati (*computer crimes*<sup>24</sup>), nonché l'urgenza di una riforma normativa e tecnica<sup>25</sup>. Le questioni sin qui evidenziate consentono di delineare un primo quadro delle criticità connesse alla gestione del dato nell'attuale scenario digitale. Tuttavia, ulteriori problematiche si manifestano con particolare intensità in presenza delle nuove architetture informatiche – quali i *data lake*, i servizi in *cloud* e i sistemi di integrazione di banche dati eterogenee – che trovano nell'impiego dell'intelligenza artificiale a fini predittivi il loro principale punto di forza, ma al tempo stesso la loro area di maggiore vulnerabilità. Le pagine che seguono approfondiscono i principali nodi critici dell'amministrazione algoritmica, esplorando i rischi derivanti dall'opacità dei modelli e dall'*automation bias*, le fragilità del consenso informato e della correlata questione del *consent bias*, le implicazioni culturali e organizzative, le ambiguità nelle responsabilità di trattamento e, infine, il tema centrale della profilazione.

### 3.1. Opacità algoritmica e motivazione

Con il diffondersi delle banche dati<sup>26</sup> e i primi scambi di informazioni tra archivi diversi – sovente in assenza di un quadro giuridico adeguato, circostanza che conferma il consueto ritardo del diritto rispetto al progresso tecnologico<sup>27</sup> – emerse con forza la richiesta di trasparenza e chiarezza nei trattamenti informativi. Gli interessati volevano sapere come sarebbero stati trattati i loro dati, accedervi e correggerli; in altre parole, nacque così la rivendicazione di un vero e proprio «diritto alla buona amministrazione informatizzata»<sup>28</sup>.

D'altronde, in un sistema giuridico fondato sul principio di legalità, la trasparenza dell'azione amministrativa non costituisce soltanto un valore

---

<sup>23</sup> L'indagine condotta a Bari ha accertato quasi 7.000 accessi non autorizzati da parte di un ex dipendente di Intesa Sanpaolo, con dati relativi a circa 3.500 clienti. Tra i soggetti spiati figuravano anche autorità politiche di vertice, come Giorgia Meloni e Ignazio La Russa, oltre a membri della magistratura e delle forze armate.

<sup>24</sup> Per una ricostruzione più puntuale cfr. M.G. Losano, *Scritti di informatica e diritto. Per una storia dell'informatica giuridica*, Milano, 2022, vol. 1.

<sup>25</sup> Sul tema della *cybersicurezza* nello specifico, vd. A. Contaldo-D. Mula (a cura di), *Cybersecurity law: disciplina italiana ed europea della sicurezza cibernetica anche alla luce delle norme tecniche*, Pisa, 2020; S. Buseti-F. M. Scanni, *Pericoli Cyber. Un'analisi dei processi di istituzionalizzazione e di Policy Making della cybersicurezza nazionale*, in S. Buseti-A. Noto-R. Romani (a cura di), *Essere Digitali. Le scienze della politica allo studio dell'ultima rivoluzione*, Teramo, 2023, 67 ss.

<sup>26</sup> Cfr. R. Ridi, *Il mondo dei documenti: cosa sono, come valutarli e organizzarli*, Bari, 2014.

<sup>27</sup> E. D'Orlando, *Politica, tecnica e scienza: il sistema delle fonti di fronte al dilemma della complessità*, in *Diritto amministrativo*, 4, 2021, spec. 723.

<sup>28</sup> Cfr. R. C. Perin-D. U. Galetta, *Il Diritto*, cit., 77 ss.

etico, ma una condizione di validità e di legittimità del potere esercitato. L'amministrazione, infatti, è vincolata a fornire motivazioni comprensibili e verificabili, in modo che il cittadino possa esercitare il diritto di difesa e il giudice possa svolgere un controllo effettivo sulla correttezza del procedimento.

L'avvento delle decisioni algoritmiche ha tuttavia introdotto un elemento di radicale discontinuità. Nello specifico, l'opacità del processo decisionale automatizzato – che può derivare da una segretezza intenzionale da parte dei soggetti pubblici o privati che sviluppano o utilizzano gli algoritmi (*intentional corporate or state secrecy*), da una diffusa incompetenza tecnica o asimmetria informativa (*technical illiteracy*) oppure dalla complessità intrinseca dei sistemi di apprendimento automatico, che operano su scala e con modalità non pienamente intellegibili nemmeno ai loro creatori (*the way algorithms operate at the scale of application*)<sup>29</sup> – incrina la logica stessa della legalità amministrativa. Essa spezza, infatti, il nesso di imputazione tra decisione e decisore:

«[...] e ciò in quanto la correlazione tra dati che vale ad estrarre la “regola” su cui la decisione si fonda viene sviluppata dalla macchina, senza quindi che tale correlazione possa essere comprensibile o nota, nemmeno a chi l'algoritmo ha creato mediante la programmazione del relativo *software*»<sup>30</sup>.

La questione, peraltro, diviene ancora più articolata se si considera la barriera aggiuntiva a questo oscurantismo che si presenta sotto forma di protezione della proprietà intellettuale, usata per giustificare la segretezza del codice sorgente o dei tipi di modelli matematici impiegati. Tale protezione, pur legittima sul piano economico, finisce per ostacolare ogni forma di ispezione, *audit* o controllo da parte degli *stakeholder* o delle autorità di regolamentazione, favorendo la creazione di quelle «*black boxes that even experts struggle to interpret, giving rise to an urgent need for transparency and interpretability to ensure accountability and trust*»<sup>31</sup>. Proprio per questo, la responsabilità del decisore si estende anche alle fasi di *procurement* e di *outsourcing* tecnologico, imponendo una diligenza rafforzata nella selezione dei fornitori e nella definizione dei contratti di servizio. In tali contratti dovrebbero essere previste clausole di cooperazione, *audit* e accesso al codice o ai *log* decisionali, in coerenza con l'art. 30 del d.lgs. 36/2023 e con l'art. 14 dell'AI Act. Ora, è doveroso riconoscere che la giurisprudenza europea e italiana conoscono bene i rischi derivanti dalla disquisita opacità.

La Corte distrettuale dell'Aia, nel caso *SyRI* (acronimo di *System Risk Indication*, il sistema algoritmico oggetto di causa), ha invalidato un sistema di sorveglianza predittiva proprio perché l'opacità sui dati e sulle regole di

---

<sup>29</sup> Così J. Burrel, *How the machine 'thinks': Understanding opacity in machine learning algorithms*, in *Big data & society*, 3, 2016, spec. 3, 4 e 5.

<sup>30</sup> G. F. Licata, *Intelligenza artificiale e contratti pubblici: problemi e prospettive*, in *CERIDAP Rivista Interdisciplinare sul diritto delle Amministrazioni Pubbliche*, 2, 2024, 30 ss.

<sup>31</sup> A. Adadi-M. Berrada, *Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)*, in *IEEE Access*, 6, 2018, spec. 52139. Sull'idea della *black box*, quale metafora dell'opacità algoritmica, vd. F. Pasquale, *The Black Box Society. The Secret Algorithms That Control Money and Information*, Cambridge, 2015.

combinazione impediva un controllo effettivo e creava un rischio sproporzionato per i diritti. Sul versante interno, il Consiglio di Stato (sez. VI, 8 aprile 2019, n. 2270) in relazione all’assegnazione del personale scolastico e dei tirocinanti, ha chiarito che segreti industriali o vincoli contrattuali non possono precludere la conoscibilità dei criteri che incidono su posizioni soggettive. Analogo principio è stato ribadito in Francia nel caso *Parcoursup*, la piattaforma digitale utilizzata per gestire le candidature post-diploma: il *Conseil d’État* (12 giugno 2020, n. 418142)<sup>32</sup> ha affermato la necessità di rendere conoscibili i criteri di calcolo in nome dell’egualianza e della verificabilità amministrativa.

Pertanto, quando la decisione pubblica si appoggia a un punteggio o a una classificazione non intelligibile, la motivazione rischia di scivolare in una sorta di circolarità tautologica, riducendosi a formule del tipo “il modello ha restituito uno score di 0,78”<sup>33</sup> e “così è deciso”<sup>34</sup>, che la rendono, in definitiva, «*not justifiable or legitimate*»<sup>35</sup>.

Si tratta di una tautologia che riflette, come è stato osservato, la pretesa del mondo digitale di «essere letteralmente tautologico», in cui «nulla può avvenire che trascenda l’immanenza dell’informazione» e in cui «la legge algoritmica, la computabilità – per usare le parole di Fredkin – diventano la legge del tutto»<sup>36</sup>. In un simile scenario, la decisione automatizzata finisce per svuotare di significato quella “riserva di umanità”<sup>37</sup> a cui il sistema dovrebbe tendere attraverso il controllo del cittadino e del giudice. Tale svuotamento trova la sua espressione più evidente in quello che la lettera-

<sup>32</sup> Sul punto cfr. L. Frouillou, *Parcoursup: quelles sélections à l’entrée dans le supérieur?*. In *Sélections, du système éducatif au marché du travail: 26 es Journées du Longitudinal. Céreq*, 2020, spec. 44-48

<sup>33</sup> I modelli di machine learning producono risultati espressi in termini di punteggi o probabilità, che rappresentano il grado di rischio o di appartenenza di un soggetto a una determinata categoria (ad esempio “alto rischio” o “potenziale recidiva”). Tuttavia, se tali score non sono accompagnati da una spiegazione dei criteri e delle variabili che li hanno generati, si riducono a meri enunciati tecnici, incapaci di assolvere alla funzione motivazionale che, nel diritto amministrativo, costituisce presidio imprescindibile di legalità e trasparenza. Sul punto, con particolare riguardo alle implicazioni in tema di trasparenza algoritmica e diritto alla spiegazione, tra gli altri, cfr. S. Wachter-B. Mittelstadt-C. Russell, *Counterfactual explanations without opening the black box: Automated decisions and the GDPR*, in *Harvard Journal of Law & Technology*, vol. 21, 2018, 841 ss.; J. Morley-A. Elhalal-F. Garcia-L. Kinsey-J. Mökander-L. Floridi, *Ethics as a service: a pragmatic operationalisation of AI ethics*, in *Minds and Machines*, 31, 2021, 239 ss.; J. Fehr-B. Citro-R. Malpani-C. Lippert-V. I. Madai, *A trustworthy AI reality-check: the lack of transparency of artificial intelligence products in healthcare*, in *Frontiers in Digital Health*, 2024, 1 ss.

<sup>34</sup> Sulla crescente delega delle decisioni ai sistemi automatizzati cfr. F. Marasà, *Intelligenza artificiale e tutela dei dati personali. Quali riflessi sulla giustizia predittiva?*, in *Osservatorio del diritto civile e commerciale*, 1, 2023, 73 ss.

<sup>35</sup> A. Barredo Arrieta-N. Díaz-Rodríguez-J. Del Ser-A. Bennetot-S. Tabik-A. Barbado-S. Garcia-S. Gil-Lopez-D. Molina-R. Benjamins-R. Chatila-F. Herrera, *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI*, in *Information Fusion*, 58, 2020, spec. 83.

<sup>36</sup> G. De Ruvo, *Dio, l’evento e l’algoritmo: il tradimento di Leibniz nell’ontologia digitale e l’etica dell’istante*, in *Segni e Comprensione XXXVI*, 103, 2022, spec. 96.

<sup>37</sup> G. Gallone, *Riserva di umanità e funzioni amministrative. Indagine sui limiti dell’automazione decisionale tra procedimento e processo*, Milano, 2023.

tura definisce *automation bias*<sup>38</sup>.

Addentrando su questa via, appare opportuno evidenziare che il cuore del problema rimane pur sempre il rapporto tra automazione e supervisione umana. In altri termini, se la retorica ufficiale insiste nel presentare l'algoritmo come un mero strumento di supporto al decisore, la prassi mostra spesso una realtà diversa: l'operatore tende a conformarsi all'*output* della macchina, specie quando manchi il tempo per un'istruttoria approfondita o quando l'autorevolezza tecnica del modello sembri incontestabile. In tali circostanze, la supervisione umana rischia di ridursi a una mera ratifica passiva<sup>39</sup>, svuotando di significato i principi di legalità, motivazione e responsabilità.

Qui si colloca il nodo critico. La presenza dell'elemento umano non basta a neutralizzare l'effetto vincolante dello *score* o della classificazione derivante dalla sua apparente neutralità e precisione. Ne deriva una progressiva delegittimazione della funzione umana, la quale rischia di ridursi – se si consente la metafora – a un ruolo meramente protocollare, privo della capacità di incidere realmente sull'esito del procedimento.

L'*automation bias* non costituisce, dunque, un incidente marginale, bensì una conseguenza sistemica dell'automazione decisionale, che si alimenta di un duplice paradosso: da un lato, l'illusione di oggettività e neutralità dell'algoritmo che genera una fiducia acritica sulla sua presunta affidabilità; dall'altro, la riduzione della discrezionalità umana a funzione accessoria, incapace di ristabilire pienamente le garanzie di motivazione, proporzionalità e responsabilità proprie del diritto amministrativo.

Un ulteriore livello di opacità si manifesta, ancor prima che nei risultati dell'elaborazione algoritmica, nella fase di selezione e accumulo dei dati. L'esigenza di sviluppare modelli sempre più performanti alimenta infatti una raccolta massiva e spesso indiscriminata di informazioni<sup>40</sup>, priva di un effettivo riscontro circa la loro pertinenza o necessità rispetto alle finalità del trattamento<sup>41</sup>. A ciò si aggiunge la tendenza a conservarle oltre il tempo strettamente necessario, nella prospettiva di riusi futuri, come l'addestramento di nuovi modelli predittivi. Ne deriva un circolo vizioso: l'espansione continua della base informativa, lungi dal migliorare la qualità dei risultati, finisce per amplificare le distorsioni già presenti e per rendere sempre più opaco il rapporto tra la finalità dichiarata e l'uso effettivo dei dati, compromettendo i principi di minimizzazione e di limitazione della conservazione e svuotando di significato il principio di finalità.

Questa dinamica mostra come la poca trasparenza algoritmica non si esau-

---

<sup>38</sup> Il concetto è ampiamente studiato nelle scienze cognitive e nell'informatica applicata. Cfr. M. L. Cummings, *Automation bias in intelligent time critical decision support systems*, in D. Harris-W-C. Li (a cura di), *Decision making in aviation*. Routledge, 2017, 289 ss.

<sup>39</sup> B. Green, *The flaws of policies requiring human oversight of government algorithms*, in *Computer Law & Security Review*, 45, 2022, 1 ss.

<sup>40</sup> Cfr. European Union Agency for Fundamental Rights, *Data quality and artificial intelligence - mitigating bias and error to protect fundamental rights*, FRA Focus, in fra.europa.eu, 2019.

<sup>41</sup> Sul rapporto tra liceità e clausola di necessità cfr. B. Ponti, *Attività amministrativa e trattamento dei dati personali: gli standard di legalità tra tutela e funzionalità*, Milano, 2023, spec. 31 ss.

risca nel momento decisionale, ma affondi le proprie radici nella costruzione stessa del sapere amministrativo. Ciò che diventa oscuro non è solo il risultato, ma il modo in cui si decide che cosa misurare, raccogliere e conservare. In tal senso, l'opacità è strutturale, riguarda il processo con cui la realtà viene tradotta in dati e trasformata in informazione utile al potere decisionale.

Essa solleva interrogativi profondi sulla compatibilità tra la logica dell'accumulo informativo e i principi di libertà e di autogoverno che fondano la democrazia. L'uso pervasivo dei sistemi di intelligenza artificiale rischia infatti di alimentare una democrazia predittiva, nella quale la partecipazione dei cittadini si riduce a una produzione costante di dati destinati ad alimentare modelli matematici e meccanismi di sorveglianza automatizzata<sup>42</sup>. In tale scenario, il richiamo del GDPR ai diritti fondamentali – dal diritto alla protezione dei dati sancito dall'art. 8 della Carta dei diritti fondamentali dell'Unione europea alla dignità intrinseca della persona (art. 1 GDPR) – non appare più sufficiente. La sola tutela formale della riservatezza, infatti, non basta a fronte di rischi più sottili ma altrettanto invasivi, come la distorsione dell'identità personale e l'ampliamento progressivo della nozione stessa di dato personale, che assume un carattere dinamico e tendenzialmente onnicomprensivo<sup>43</sup>.

In questa prospettiva, la trasparenza e la spiegabilità dei modelli non appaiono più come mere esigenze tecniche, ma come condizioni essenziali per preservare l'*accountability* e la fiducia pubblica nel processo decisionale automatizzato.

Certo, fondare la soluzione del problema su questi principi generale potrebbe rilevarsi insufficiente. Non può, infatti, essere trascurata anche la dimensione culturale del problema. L'amministrazione predittiva richiede nuove competenze non soltanto tecniche, ma anche analitiche e valutative. È necessaria, pertanto, un'adeguata alfabetizzazione ai dati da parte dei funzionari pubblici, una crescente capacità di leggere indicatori e risultati senza assumerli come verità oggettive, nonché una piena consapevolezza dei rischi di *bias*, *overfitting*, *drift* e *fairness*. Senza questo salto intellettuale, la promessa di un effettivo *human in the loop*<sup>44</sup> rimane nominale, e la retorica dell'innovazione rischia di prevalere sulla reale capacità di vigilanza.

---

<sup>42</sup> L.G. Sciannella, *Intelligenza artificiale, politica e democrazia: Artificial Intelligence, Politics and Democracy*, in *Diritto Pubblico Comparato ed Europeo online*, 1, 2022, spec. 337.

<sup>43</sup> In dottrina si è evidenziato come la distinzione tra dato personale e dato non personale stia progressivamente perdendo chiarezza, poiché le pratiche di *data mining* e di profilazione rendono difficile, se non impossibile, una separazione netta tra le due categorie. La commistione di elementi identificativi e non identificativi nei *dataset* eterogenei e dinamici conduce, di fatto, a un'erosione della tradizionale dicotomia. Sul punto, vd. S. Torregiani, *La circolazione dei dati secondo l'ordinamento giuridico europeo. Il rischio dell'ipertrofia normativa*, in *Rivista italiana di informatica e diritto*, 1, 2021, spec. 48 ss.; C. Bergonzini, "Prova a prendermi". *Ecosistema digitale e consapevolezza degli utenti: uno spazio per la regolazione nazionale?*, in G. Di Cosimo (a cura di), *Processi democratici e tecnologie digitali*, Torino, 2023, spec. 97. La sfumata (*blurred*) distinzione tra dati personali e non personali è stata constatata anche da parte dell'European Parliamentary Research Service (EPRS), così M. D. M. Negreiro Achiaga, *Free flow of non-personal data in the European Union*, in *Briefing, EU Legislation in Progress*, 2019, 10.

<sup>44</sup> D. Martire, *Human in the loop. L'essere umano come fattore condizionante della—o condizionato dalla—intelligenza artificiale*, in *Rivista italiana di informatica e diritto*, 2, 2024, 1 ss.

L'analisi sin qui condotta mostra come l'opacità algoritmica non incida soltanto sul come si decide, ma anche sul modo in cui le decisioni vengono percepite e accettate da chi ne è destinatario. La medesima logica di deresponsabilizzazione che condiziona il decisore si riflette, infatti, sull'interessato, erodendo la consapevolezza e la libertà del consenso. L'automazione non altera solo il processo decisionale interno all'amministrazione, ma anche il modo in cui il cittadino "acconsente" ai trattamenti che lo riguardano, spesso senza comprenderne la portata né le implicazioni pratiche. Diventa allora necessario interrogarsi sul valore effettivo del consenso nell'era predittiva: se esso rappresenti ancora uno strumento di autodeterminazione o se, al contrario, stia divenendo una mera formalità, priva di reale capacità di controllo. A questo tema sarà dedicato il paragrafo successivo.

### 3.2. Il consenso come base fragile

Non meno problematico è il riflesso che la mancanza di trasparenza esercita sulla nozione di consenso informato, in particolare quando il trattamento dei dati avviene su base consensuale. In questo contesto, il diritto alla protezione dei dati personali non può essere più considerato solo una garanzia della riservatezza individuale, ma si configura come presidio fondamentale all'effettivo esercizio delle libertà nel contesto digitale<sup>45</sup>. Eppure, nemmeno il consenso del cittadino rappresenta una garanzia effettiva. La lettera del GDPR, infatti, consente il trattamento automatizzato fondato sul consenso esplicito dell'interessato (art. 22, par. 2, lett. c)), ma, anche quando tale consenso viene formalmente richiesto, non può qualificarsi come realmente libero e informato. Nessun utente è in grado di comprendere *ex ante* quali combinazioni di dati alimentino i modelli, quali risulti futuri siano possibili e quali conseguenze pratiche possano derivarne. Ne consegue che, nei sistemi predittivi, la nozione stessa di consenso rischia di ridursi a una mera finzione giuridica, scarsamente utile alla protezione sostanziale della persona e, anzi, idonea a ostacolare una corretta allocazione della responsabilità in capo al titolare del trattamento. Insomma, in questo senso, si può affermare che il consenso, pur formalmente acquisito, risulta sostanzialmente viziato da asimmetria informativa e asimmetria cognitiva: la volontà si forma su presupposti ignoti e l'autodeterminazione si svuota di contenuto. Com'è stato osservato:

«rimane oscuro come e a quali condizioni il funzionario umano possa legittimamente discostarsi dalle risultanze istruttorie algoritmiche, e come viceversa debba attivarsi per rilevare condotte atipiche e diverse o introdurre elementi di valutazione ulteriori che possono capovolgere il contenuto stesso della decisione finale»<sup>46</sup>.

---

<sup>45</sup> S. De Luca, *Big Data, libertà e democrazia: tra utopia e distopia*, in S. Busetti-A. Noto-R. Romani (a cura di), *Essere Digitali*, cit., 27 ss.

<sup>46</sup> G. Avanzini, *Intelligenza artificiale, machine learning e istruttoria procedimentale: vantaggi, limiti ed esigenze di una specifica data governance*, in *Intelligenza artificiale e diritto: una rivoluzione?*, in *Quaderni ASTRID*, vol. 2, Bologna, 2022, spec. 92.

Sul piano dogmatico, ciò comporta un duplice rischio. Da un lato, l'amministrazione non è più in grado di valutare in modo trasparente se il mezzo sia necessario e idoneo rispetto al fine, alterando così il principio di proporzionalità del trattamento. Dall'altro, la nozione di consenso come fondamento di liceità si svuota di significato, riducendosi a mera formalità priva di una effettiva capacità di garanzia. Ne risulta un meccanismo che deresponsabilizza la pubblica amministrazione più che garantire l'autodeterminazione individuale. La libertà del consenso evapora nella complessità epistemica del sistema.

A ben vedere, la problematica non è nuova ma ripropone, in chiave digitale, il tema classico dell'eterodeterminazione della volontà. Tuttavia, nel contesto algoritmico, l'eterodeterminazione non deriva più da un condizionamento esterno, bensì da una struttura epistemica opaca che confina il cittadino in una posizione di inconsapevolezza strutturale. La decisione amministrativa perde così il suo carattere dialogico<sup>47</sup>, trasformandosi in un'interazione unidirezionale in cui la macchina determina i confini della partecipazione e della conoscenza.

Le criticità del consenso non si limitano però al piano soggettivo, ma si riflettono anche su quello sistemico.

Il cosiddetto *consent bias* – che si manifesta quando le analisi si basano solo sui soggetti che hanno prestato consenso al trattamento – introduce distorsioni che compromettono la rappresentatività dei campioni, generando discriminazioni di fatto. Tale rischio è particolarmente evidente in settori sensibili, come la sanità o la ricerca biomedica, dove motivi etici o personali riducono la disponibilità a condividere dati. In tali casi, quindi, il rispetto formale dell'autodeterminazione informativa può tradursi in risultati fuorvianti e potenzialmente iniqui.

Inoltre, il caso dell'applicazione *Weople*, che proponeva un modello di "economia dei dati" in cui gli utenti cedono informazioni personali in cambio di una remunerazione, mostra ulteriori limiti<sup>48</sup>. Il Garante per la protezione dei dati personali, già nel 2019, segnalò al Comitato europeo per la protezione dei dati (d'ora in avanti, EDPB) i rischi di tale pratica: da un lato, possibili duplicazioni e violazioni dei principi di integrità e minimizzazione; dall'altro, la trasformazione del consenso in merce negoziabile, in contrasto con la sua natura di atto libero e informato ai sensi dell'art. 7 GDPR. La stessa Autorità ha poi richiamato l'esigenza di vigilare sulla legittimità delle richieste di portabilità dei dati (art. 20 GDPR), sottolineando che solo i dati "forniti" dall'interessato e trattati con mezzi automatizzati possono essere trasferiti, non quelli derivati o inferiti<sup>49</sup>.

In definitiva, il consenso, lungi dal rappresentare una garanzia di libertà, si rivela oggi una base fragile. Esso appare sempre più come uno strumento

---

<sup>47</sup> In dottrina si parla di "diritto all'interazione digitale" tra pubblica amministrazione e cittadini: cfr. E. Belisario, *La "nuova" Pubblica Amministrazione digitale*, Rimini, 2009.

<sup>48</sup> Sul caso *Weople*, cfr. F. Bravo, *Intermediazione di dati personali e servizi di data sharing dal GDPR al Data Governance Act, in Contratto e impresa Europa*, 1, 2021, spec. 216-221.

<sup>49</sup> Art. 29 WORKING PARTY, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, adopted on 3 October 2017, revised on 6 February 2018, 19.

formale, piuttosto che sostanziale, e finisce per operare più come meccanismo di deresponsabilizzazione che come effettiva tutela dell'interessato. La crisi del consenso, resa particolarmente evidente dall'opacità algoritmica, impone dunque di ripensare le modalità di legittimazione dei trattamenti automatizzati e di rafforzare i meccanismi di imputazione della responsabilità.

Questa fragilità non costituisce soltanto il sintomo di una crisi della libertà informativa. Essa rappresenta, più in profondità, il preludio a una crisi della responsabilità. Quando l'autodeterminazione dell'interessato risulta compromessa dall'opacità dei processi decisionali automatizzati, anche l'imputazione delle decisioni e dei loro effetti diventa incerta.

In tale prospettiva, il tema della responsabilità amministrativa emerge come il naturale sviluppo della riflessione, ponendo l'interrogativo su come attribuire, nel contesto automatizzato, la titolarità del potere e la conseguente rendicontabilità dell'azione pubblica.

### **3.3. Responsabilità amministrativa e *accountability* progettuale**

La crisi del consenso come strumento di legittimazione dei trattamenti automatizzati, come appena anticipato, apre inevitabilmente la questione della responsabilità amministrativa. D'altra parte, se il cittadino non è più in grado di comprendere i processi decisionali fondati su algoritmi opachi, il problema non è solo garantire l'autodeterminazione informativa, ma assicurare che l'esercizio del potere resti imputabile e controllabile. In altri termini, la fragilità del consenso si traduce in una fragilità della responsabilità: là dove viene meno la comprensibilità della decisione, si affievolisce anche la possibilità di individuare chi debba risponderne.

È in questo passaggio che il tema dell'*accountability* si intreccia con quello della responsabilità amministrativa. Entrambi esprimono, seppur su piani differenti, l'esigenza che il potere resti tracciabile e giustificabile: l'*accountability* sul versante organizzativo e tecnico, la responsabilità su quello giuridico e personale. In questa prospettiva, il GDPR offre una prima traduzione normativa del principio, imponendo al titolare del trattamento di dimostrare la conformità delle operazioni ai criteri di liceità, correttezza e trasparenza. Tale obbligo, tuttavia, non esaurisce la questione sul piano sostanziale; non basta che il procedimento rispetti formalmente le prescrizioni regolamentari, è necessario che sia anche comprensibile, motivato e controllabile nella sua logica decisionale.

In questa prospettiva, nel contesto amministrativo, l'*accountability* si lega con il principio costituzionale di buon andamento (art. 97 Cost.) e con l'obbligo di motivazione, assumendo un significato più denso rispetto a quello meramente regolatorio. Essa diventa la misura della trasparenza del potere pubblico e della sua capacità di rendere conto delle scelte algoritmiche adottate.

Ciò posto, se la decisione è il frutto di un processo algoritmico non spiegabile, l'amministrazione si trova in una condizione di irresponsabilità strutturale, poiché non è più in grado di ricostruire la catena causale che

conduce dall'*input* al risultato. In questa situazione, il potere pubblico rischia di trasformarsi in un potere “anonimo”, in cui la decisione appare come un fatto tecnico piuttosto che come un atto imputabile. L'effetto, in termini di garanzie, è dirompente giacché si dissolve la possibilità stessa di individuare il responsabile, di motivare la decisione e di consentire al cittadino di contestarla efficacemente.

Da qui nasce l'esigenza di ripensare la responsabilità amministrativa nell'era della predizione: essa non può più limitarsi al piano soggettivo, ma deve essere ricostruita come responsabilità progettuale (o multilivello), che abbraccia l'intero ciclo di vita dell'algoritmo: dalla progettazione all'addestramento, fino all'utilizzo operativo. Non si tratta più soltanto di reagire a un danno *ex post*, ma di garantire, *ex ante*, che la tecnologia sia progettata e utilizzata secondo criteri di affidabilità, trasparenza e proporzionalità. Pertanto, essa coinvolge non solo i funzionari pubblici, ma anche i fornitori, le stazioni appaltanti e i soggetti esterni, secondo logiche di prevenzione e di governance ispirate all'art. 30 del d.lgs. 36/2023 e al modello di supervisione previsto dall'AI Act (art. 14)<sup>50</sup>.

La logica sottostante è quella del principio di continuità della responsabilità, radicato nell'art. 28 Cost., secondo cui il potere decisionale, anche quando mediato da strumenti tecnologici, non può mai essere disgiunto dall'obbligo di comprensione, motivazione e controllo. Ne consegue che, anche in presenza di un sistema algoritmico, il funzionario è tenuto a rispondere delle conseguenze derivanti dall'uso di strumenti di cui non comprende pienamente il funzionamento, a meno che non abbia dimostrato di aver adottato tutte le misure necessarie per verificarne l'affidabilità e la correttezza.

L'uso di algoritmi predittivi non può essere considerato una delega neutra di funzioni tecniche; esso comporta, invece, l'assunzione di una responsabilità complessa, che si estende alla scelta, alla configurazione e alla validazione del modello impiegato. In questa prospettiva, la responsabilità del decisore pubblico assume una dimensione *ex ante*, di prevenzione e controllo, e non più soltanto *ex post*, di sanzione o riparazione del danno. L'applicazione coerente di tale principio esige che l'amministrazione dimostri non solo di aver effettuato una corretta valutazione d'impatto sulla protezione dei dati (DPIA, ne riparleremo più avanti), ma anche di aver predisposto procedure interne di *audit* e supervisione. L'assenza di tali strumenti può integrare un vizio di *culpa in vigilando* o di *culpa in eligendo*, configurando una responsabilità amministrativa diretta in caso di danno ingiusto derivante dall'uso improprio del modello.

Sul versante europeo, la giurisprudenza della Corte di giustizia ha contribuito a definire questa esigenza di controllo effettivo. Nella causa *H.K. v. Prokuratuur* (C-746/18, 2021), la Corte ha stabilito che l'accesso ai dati di traffico e di localizzazione incide gravemente sui diritti fondamentali e può essere autorizzato solo previa verifica preventiva da parte di un giudice o di un'autorità indipendente e imparziale<sup>51</sup>. Sebbene la pronuncia maturi in

---

<sup>50</sup> Cfr. M. Barberio, *L'uso dell'intelligenza artificiale nell'art. 30 del d.lgs. 36/2023 alla prova dell'AI Act dell'Unione europea*, in *Rivista italiana di informatica e diritto*, 2, 2023, 253 ss.

<sup>51</sup> Sul punto vd. I. Revolidis, *HK v Prokuratuur: On Balancing Crime Investigation and Data Protection*, in *European Data Protection Law Review*, 6, 2020, 319 ss.; S. Rovelli, *Case*

ambito penale, la ratio del controllo effettivo (e non meramente simbolico) è estensibile anche al procedimento amministrativo automatizzato, in virtù del principio comune di proporzionalità e di effettiva tutela giurisdizionale.

L'AI Act recepisce questa impostazione introducendo un obbligo generale di supervisione umana effettiva per i sistemi di intelligenza artificiale ad alto rischio (art. 14). Tali sistemi devono essere progettati per consentire agli operatori di comprendere limiti e capacità del modello e restare consapevoli della sua influenza sul processo decisionale. La responsabilità si sposta così dal livello meramente giuridico a quello organizzativo e tecnico, imponendo la creazione di filtri istituzionali che assicurino tracciabilità e revisione delle decisioni automatizzate.

Quanto detto ci impone una riflessione sul significato della colpa algoritmica. Essa non coincide con l'errore del codice, ma con la mancata vigilanza, previsione o comprensione delle implicazioni derivanti dal suo uso. L'algoritmo, come strumento di esercizio del potere, non è mai un soggetto autonomo, bensì un mezzo; sicché, la sua imprevedibilità non elimina la colpa, ma la trasforma in colpa per difetto di vigilanza.

In questo senso, la responsabilità del decisore algoritmico si colloca tra la responsabilità da atto e la responsabilità da organizzazione, poiché ciò che viene meno non è soltanto la correttezza dell'atto singolo, ma la capacità del sistema amministrativo di garantire un esercizio del potere conforme a diritto. Si direbbe che la vera sfida non sia adattare l'algoritmo al diritto, ma adattare il diritto a controllare l'algoritmo, nel solco della dottrina che propone di reinterpretare la legalità in chiave di "technological due process"<sup>52</sup>. Si comprende allora la necessità di ricercare un nuovo equilibrio tra fiducia tecnica e controllo giuridico: la sfida, in altri termini, consiste nel ricondurre la discrezionalità algoritmica entro i confini della discrezionalità amministrativa. L'amministrazione, infatti, non può limitarsi a utilizzare modelli predittivi: deve comprenderli, valutarli e saperne rendere conto. D'altronde, sembrerebbe essere lo stesso obiettivo del GDPR quello di «anteporre l'uomo alla macchina algoritmica, o se si vuole, porre la macchina in modalità servente all'uomo»<sup>53</sup>, obiettivo che, tuttavia, resta in larga parte disatteso.

La riflessione sulla responsabilità amministrativa e sull'*accountability* progettuale segna dunque un passaggio decisivo nella comprensione del rapporto tra amministrazione e tecnologia. Quando trasparenza e motivazione vengono meno, si apre la strada a un potere algoritmico opaco, potenzialmente lesivo dell'eguaglianza e della fiducia pubblica.

È proprio su questo terreno che si colloca la successiva analisi: quella dei

---

*Prokuraatur: proportionality and the independence of authorities in data retention*, in *European Papers-A Journal on Law and Integration*, 1, 2021, 199 ss.; G. Formici, *La sentenza HK c. Prokuraatur e il difficile dialogo tra CGUE e Stati membri in materia di conservazione e accesso ai metadati per finalità securitarie: spunti di riflessione su una questione vecchia ma ancora irrisolta*, in *Quaderni di SIDIBlog*, Napoli, 2021, 231 ss.

<sup>52</sup> M. C. Cavallaro-G. Smorto, *Decisione pubblica e responsabilità dell'amministrazione nella società dell'algoritmo*, in *Federalismi.it*, 16, 2019, spec. 17.

<sup>53</sup> S. Sassi, *Gli algoritmi nelle decisioni pubbliche tra trasparenza e responsabilità*, in *Analisi Giuridica dell'Economia*, 1, 2019, spec. 114.

*bias* algoritmici e della profilazione predittiva, dove il problema non è più soltanto chi risponde della decisione, ma come le decisioni stesse possono produrre discriminazioni sistemiche, in contrasto con il principio costituzionale di eguaglianza.

### 3.4. Profilazione, *bias* e principio di eguaglianza

L'art. 4, par. 4, del GDPR definisce la profilazione come qualsiasi forma di trattamento automatizzato di dati personali finalizzata a valutare aspetti personali relativi a una persona fisica.

Nei sistemi predittivi essa costituisce, pertanto, il meccanismo tecnico che trasforma dati in inferenze e le inferenze in decisioni, classificando individui e gruppi in base a somiglianze statistiche e incidendo, così, sulla distribuzione di opportunità e oneri.

Ogni operazione di profilazione implica una scelta: quali caratteristiche contano e quali possono essere trascurate, quali variabili siano predittive e quali irrilevanti. In questo senso, la selezione delle variabili rappresenta un atto di discrezionalità pubblica, che deve restare tracciabile, motivata e sottoposta a controllo. Delegare tali decisioni al modello significa spostare il baricentro della responsabilità democratica verso un dominio opaco di scelte algoritmiche.

Nel quadro dell'amministrazione predittiva, la profilazione rappresenta, dunque, il punto di massima tensione tra efficienza algoritmica e principio di eguaglianza<sup>54</sup>.

Come è stato osservato<sup>55</sup>, infatti, l'idea di una *fairness* (equità) automatica è logicamente incoerente perché la giustizia distributiva, implicando un giudizio di valore, non può essere calcolata. Questa tensione non è soltanto tecnica ma giuridica e ontologica.

Il diritto dell'eguaglianza si fonda su comparazioni normative, su giudizi di rilevanza e di giustificazione delle differenze che dipendono dal fine pubblico perseguito. L'algoritmo, viceversa, ragiona per correlazioni e non per cause, e proprio per questo tende per sua natura a replicare e cristallizzare *pattern* di disuguaglianze sociali preesistenti. Le decisioni automatizzate, addestrate su archivi storici, finiscono così per creare nuovi *cluster* (*social sorting*<sup>56</sup>) fondati su comportamenti solo apparentemente neutri – come stili di vita o abitudini alimentari – generando discriminazioni personalizzate che il diritto dovrebbe contrastare, non legittimare. La discriminazione assume allora la forma di una ingiustizia distributiva automatizzata, che riduce l'eguaglianza sostanziale a una media statistica.

Ciò dimostra che i dati non sono meri strumenti di conoscenza, ma veicoli di costruzione sociale della realtà. E poiché ogni costruzione del reale è anche un atto di potere, il diritto dei dati non appartiene alla sfera della tec-

---

<sup>54</sup> Cfr. G. Resta, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del diritto*, 2, 2019, 199 ss.

<sup>55</sup> S. Wachter-B. Mittelstadt-C. Russell, *Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI*, in *Computer Law & Security Review*, 41, 2021, 1 ss.

<sup>56</sup> Il sociologo David Lyon descrive efficacemente questo fenomeno: D. Lyon, *Surveillance as social sorting: Privacy, risk and automated discrimination*, Londra, 2005.

nica, ma a quella inevitabilmente politica dei rapporti che essa contribuisce a istituzionalizzare<sup>57</sup>.

La *fairness* computazionale, pertanto, può offrire solo indicatori descrittivi di disuguaglianza, ma non può sostituire il giudizio giuridico, che resta un atto di ponderazione valoriale. L'idea di un ragionamento predittivo neutrale diviene, così, un'illusione, un mito tecnico e giuridico<sup>58</sup> che sbiadisce il carattere normativo delle scelte algoritmiche e, in questa prospettiva, la profilazione diventa banco di prova della capacità del principio di eguaglianza sostanziale di agire anche nel dominio dei dati.

Sotto il profilo dell'eguaglianza, è emblematico il caso COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) negli Stati Uniti, in cui un algoritmo di *risk assessment* classificava sistematicamente come "ad alto rischio" soggetti afroamericani non recidivi. Nel caso *de qua*, la discriminazione deriva non da un'intenzione, ma da correlazioni spurie e variabili *proxy*. In termini giuridici, si tratta del cosiddetto *disparate impact*<sup>59</sup>, una forma di discriminazione indiretta prodotta da regole apparentemente neutrali ma dagli effetti diseguali.

In questo quadro, sia il *disparate impact* statunitense sia la discriminazione indiretta del diritto europeo si fondano su una logica analoga: non punire l'intenzione soggettiva, ma valutare *ex post* la ragionevolezza del criterio utilizzato e i suoi effetti differenziali.

Ciò conferma la tesi che abbiamo via via delineato: l'eguaglianza non può essere tradotta in un'unica metrica di *fairness*, poiché la giustificazione delle differenze è sempre situata e dipendente dal contesto istituzionale e dal fine pubblico perseguito<sup>60</sup>.

Il principio di eguaglianza, però, non è solo un limite, ma una condizione di legittimità della predizione amministrativa.

L'eguaglianza sostanziale, nella lettura offerta dall'art. 3, c. 2, Cost., impone di interrogare le metriche tecniche alla luce dei valori costituzionali. Ogni soglia, ogni funzione di perdita, ogni parametro di accuratezza incorpora un giudizio di valore che deve essere giustificato in termini di proporzionalità e ragionevolezza. In questo senso, la *fairness* computazionale è uno strumento di evidenza, non criterio di giudizio.

Da qui discende l'esigenza di un diritto dell'algoritmo capace di integrare qualità dei dati e rappresentatività sociale. La qualità, in questa prospettiva,

---

<sup>57</sup> M. Hildebrandt-A. De Bois, *Law for Computer Scientists*, in M. Chetouani-V. Dignum-P. Lukowicz-C. Sierra (a cura di), *In Human-centered artificial intelligence: Advanced lectures*, Svizzera, 2023, spec. 261.

<sup>58</sup> *Ex multis* vd. M. Airoidi-D. Gambetta, *Sul mito della neutralità algoritmica*, in A. Martella-E. Campo-L. Ciccarese (a cura di), *Gli algoritmi come costruzione sociale*, in *The Lab's Quarterly*, 4, 2018, 25-46; M. De Simone, *Oltre il tipping point dell'Intelligenza Artificiale*, in *For: rivista per la formazione*, 3, 2023, spec. 29.

<sup>59</sup> Per un approfondimento sulle due dottrine giuridiche dominanti "disparate treatment" e "disparate impact", e sulle più recenti proposte di ulteriori categorie, come il "disparate mistreatment" (disparità nei tassi di errore tra gruppi di individui), si veda D. E. Moore, *Disparate treatment versus disparate impact: A distinction without a difference*, in *Syracuse Law Review*, 41, 1990, 965 ss.; M. B. Zafar-I. Valera-M. Gomez Rodriguez-K. P. Gummedi, *Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment*, in *Proceedings of the 26th international conference on world wide web*, 2017, 1171 ss.

<sup>60</sup> S. Wachter-B. Mittelstadt-C. Russell, *Why fairness cannot be automated*, cit., spec. 22-28.

non coincide con l'accuratezza statistica, ma comprende la neutralità, la completezza e la capacità del *dataset* di rappresentare in modo proporzionato i gruppi interessati.

L'AI Act recepisce parzialmente questo paradigma, imponendo per i sistemi ad alto rischio obblighi di qualità e di *audit* periodico, ma il suo approccio resta prevalentemente procedurale. *Recta via*, manca un vero principio di "substantive equality by design", una progettazione orientata agli effetti distributivi delle decisioni automatizzate<sup>61</sup>.

Una *governance* dell'algoritmo coerente con l'art. 3, c. 2, Cost. dovrebbe invece prevedere strumenti di valutazione dell'impatto sociale, in grado di misurare la capacità del sistema di promuovere l'inclusione e di prevenire discriminazioni indirette.

La difficoltà di eliminare *ex ante* o *ex post* i *bias* deriva dalla complessità tecnica dell'"unlearning", ossia dei procedimenti di disapprendimento dell'errore<sup>62</sup>. In tale contesto, la *protezione by design* rappresenta un approccio promettente: integrare misure di sicurezza e di equità sin dalla progettazione dei trattamenti consente di limitare gli errori di selezione<sup>63</sup> e garantire tutele generalizzate, indipendenti dal consenso individuale.

Le *Privacy Enhancing Techniques* (PETs), come la *Secure Multi-Party Computation*, i *commitment* crittografici, la *Homomorphic Encryption* e i metodi di *intersection*, costituiscono strumenti tecnologici maturi per proteggere i dati sensibili lungo tutto il ciclo di trattamento<sup>64</sup>. Le PETs rafforzano tanto la *input privacy* (confidenzialità dei dati di partenza) quanto la *output privacy* (protezione contro la rivelazione di informazioni nei risultati), riducendo il rischio di accessi indebiti e mitigando i *bias* sistemici.

Tuttavia, tali garanzie tecniche non bastano.

Devono essere accompagnate da un impegno giuridico e culturale nella supervisione umana. L'eguaglianza non si tutela solo con norme, ma con la formazione e la responsabilizzazione di chi utilizza gli algoritmi.

A ben vedere, infatti, la giurisprudenza amministrativa europea già si muove in questa direzione. Dalle decisioni italiane sull'algoritmo ministeriale per l'assegnazione dei docenti (Tar Lazio, sez. II-Bis, 7 agosto 2017, n. 9230) a quelle francesi su *Parcoursup* e spagnole in materia di distribuzione

---

<sup>61</sup> Cfr. S. Fredman, *Substantive equality revisited*, in *International Journal of Constitutional Law*, 14, 2016, 712 ss. L'autrice propone un approccio quadrimensionale alla *substantive equality* (riduzione dello svantaggio; contrasto a stigma e stereotipi; rafforzamento di voce/partecipazione; accomodamento della differenza e cambiamento strutturale), che qui assumiamo come base teorica per un principio di "substantive equality by design".

<sup>62</sup> U. Ruffolo-A. Amidei, *Diritto dell'intelligenza artificiale: Vol. 1: Responsabilità. Contratto. Regolazione. Veicoli autonomi*, Roma, 2024, spec. 21.

<sup>63</sup> L'eterogeneità non efficacemente governata può generare "allucinazioni statistiche", che occultano relazioni causali e compromettono la qualità delle scelte pubbliche. Per uno studio del fenomeno, soprattutto nei modelli linguistici di grandi dimensioni (LLM, *large language models*) si veda: A. Bruno-P. L. Mazzeo-A. Chetouani-M. Tliba-M. A. Kerkouri, *Insights into classifying and mitigating LLMs' hallucinations*, arXiv preprint arXiv:2311.08117, 2023; F. Wang, *LightHouse: A survey of AGI hallucination*, arXiv preprint arXiv:2401.06792, 2024; C. Novelli-F. Casolari-P. Hacker-G. Spedicato-L. Floridi, *Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity*, arXiv preprint arXiv:2401.07348, 2024.

<sup>64</sup> E. Belissario-G. Cassano, *Intelligenza artificiale per la pubblica amministrazione*, in *Diritto di Internet e tutele dei nuovi diritti*, Pisa, 2023, spec. 206 ss.

assistenziale, emerge un principio comune: la legalità dell'automazione è subordinata alla spiegabilità sostanziale del processo, che deve consentire al giudice e all'interessato di ricostruire la catena causale tra dati, regole e decisione. In questa chiave, la trasparenza non è un accessorio tecnico, ma l'antitesi della "discriminazione silenziosa".

La *fairness* computazionale, dunque, può avere un ruolo importante come strumento di evidenza, non di giudizio. Modelli come il *Conditional Demographic Disparity*<sup>65</sup> possono fornire indicatori statistici di potenziali disparità, ma spetta al diritto (e alla giurisdizione) valutarne la legittimità alla luce dei principi costituzionali.

L'eguaglianza, dunque, non è automatizzabile, è un processo deliberativo, umano e pubblico, che richiede motivazione e confronto.

In conclusione, l'eguaglianza è la bussola normativa dell'amministrazione predittiva. Non frena l'innovazione, ma la orienta, imponendo che la predizione resti compatibile con la dignità della persona e che la tecnologia rimanga strumento, non soggetto, del potere pubblico. Solo una progettazione fondata su *fairness*, *accountability* e *human oversight* può impedire che l'efficienza si trasformi in arbitrio statistico.

Su questo terreno si innesta il quadro europeo in materia di decisioni automatizzate. Il punto di convergenza tra eguaglianza e responsabilità algoritmica è rappresentato dal GDPR, cardine dell'intero sistema di tutela. È proprio al GDPR – e in particolare al suo art. 22 – che occorre ora volgere l'attenzione, per comprendere come il diritto europeo cerchi di bilanciare innovazione tecnologica e controllo umano, evitando che la predizione si traduca in una forma di spossessamento della libertà individuale.

#### **4. L'art. 22 GDPR: decisioni automatizzate e diritto alla spiegazione**

Dopo aver esaminato le tensioni tra predizione, responsabilità e uguaglianza, il passo successivo è interrogarsi sul modo in cui il diritto europeo tenta di disciplinare queste dinamiche.

Il GDPR costituisce la risposta più compiuta a tale esigenza, e in particolare il suo art. 22, dedicato alle decisioni automatizzate, rappresenta il fulcro della tutela contro l'eccesso di automazione<sup>66</sup>. Tale disposizione è stata sin dall'inizio una delle più dibattute e controverse, poiché tocca un nodo essenziale: fino a che punto è lecito affidare a un algoritmo, piuttosto che

---

<sup>65</sup> S. Wachter-B. Mittelstadt-C. Russell, *Why fairness cannot be automated*, cit., 24 ss.

<sup>66</sup> In Italia, il principale riferimento normativo in materia di protezione dei dati personali è il d.lgs. 196/2003 (Codice della Privacy), che — pur modificato più volte, da ultimo con il d.lgs. 101/2018 — continua a costituire il quadro di riferimento per il trattamento dei dati da parte delle pubbliche amministrazioni. Le successive riforme hanno progressivamente ampliato i poteri di trattamento e condivisione dei dati da parte delle PA, in nome di una governance più efficiente e interconnessa. Tuttavia, parte della dottrina (in particolare cfr. G. Carullo, *Dati personali e fini pubblici: dubbi di compatibilità europea del Codice Privacy*, in *Rivista Interdisciplinare sul diritto delle Amministrazioni Pubbliche*, 3, 2024, 37 ss.) osserva che tale estensione abbia finito per indebolire le garanzie originarie a tutela dell'individuo e per limitare l'efficacia del controllo europeo, lasciando agli Stati membri un'eccessiva autonomia nel disciplinare la materia.

a un decisore umano, la valutazione di interessi che incidono significativamente sulla sfera giuridica e personale dei cittadini? La questione non è soltanto tecnica, ma investe direttamente la legittimazione dell'azione pubblica e la sua compatibilità con i principi costituzionali<sup>67</sup> e sovranazionali che tutelano la dignità, la libertà e l'eguaglianza.

Il par. 1 della norma *de qua* riconosce, infatti, all'individuo il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona. Si tratta, dunque, non di una semplice facoltà azionabile a richiesta, ma di un divieto generale imposto al titolare del trattamento, che opera *ex lege* e *by default*, senza necessità di una previa opposizione da parte dell'interessato.

Questa architettura, apparentemente lineare nella sua sintetica formulazione giuridica, cela però nodi interpretativi complessi.

Il primo riguarda la definizione di decisione unicamente automatizzata. Nella prassi amministrativa e organizzativa, raramente un processo decisionale si presenta come integralmente delegato a un algoritmo, senza alcuna forma di controllo umano. Più spesso si tratta di procedure ibride, in cui l'algoritmo produce uno *score*, un *alert* o una raccomandazione che poi viene confermata, più o meno criticamente, da un funzionario. Ciò chiarito, se si seguisse un'interpretazione letterale, si potrebbe ritenere che il semplice tocco umano sia sufficiente per sottrarre l'intera procedura all'applicazione dell'art. 22, con il rischio di svuotarne la portata. Proprio per evitare tale esito, l'EDPB, nelle sue Linee guida del 2018 su *automated decision-making* e profilazione, ha sottolineato che «per aversi un coinvolgimento umano, il titolare del trattamento deve garantire che qualsiasi controllo della decisione sia significativo e non costituisca un semplice gesto simbolico»<sup>68</sup>.

A ben vedere, però, questa osservazione non scioglie il dubbio iniziale: sposta semplicemente la questione su un nuovo piano. Non si tratta più di stabilire quando una decisione sia automatizzata, ma quando il controllo umano possa dirsi realmente significativo e non meramente simbolico.

Come si è già visto a proposito dell'*automation bias* e del ruolo del funzionario, la presenza umana non è di per sé garanzia di consapevolezza o autonomia decisionale. Se il soggetto incaricato non dispone delle competenze necessarie per comprendere la logica del modello, del tempo per esaminare criticamente l'*output*, o dell'autorità per modificarlo, il suo intervento si riduce alla già paventata ratifica automatica.

Difatti, parte della dottrina più sensibile<sup>69</sup> ha sottolineato come questa su-

---

<sup>67</sup> In materia, oltre ai già citati Autori, si rimanda a A. Vedder-L. Naudts, *Accountability for the use of algorithms in a big data environment*, in *International Review of Law, Computers & Technology*, 31, 2017, 206 ss.

<sup>68</sup> EDPB, *Article 29 data protection working party. Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679*. 17/IT WP 251 rev.01, 2018, spec. 23.

<sup>69</sup> Cfr. C. Colapietro-A. Moretti, *L'Intelligenza Artificiale nel dettato costituzionale: opportunità, incertezze e tutela dei dati personali*, in *BioLaw Journal-Rivista di BioDiritto*, 3, 2020, spec. 382-383.

pervisione apparente sia incompatibile con la *ratio* della norma, che mira invece a preservare uno spazio effettivo di valutazione critica umana come condizione essenziale per la tutela della persona. Ne consegue che la distinzione tra decisione automatizzata e decisione assistita dall'uomo diventa labile e che il vero criterio non può essere formale, ma sostanziale, richiedendo di verificare la qualità e l'effettività dell'intervento umano.

Le difficoltà interpretative in ordine alla possibilità di esercitare un controllo significativo da parte dell'uomo trovano un'eco, *mutatis mutandis*, anche nel secondo snodo cruciale dell'art. 22: la portata della clausola sugli "effetti giuridici" o "significativi". Mentre i primi sono relativamente facili da individuare – si pensi ad un atto che attribuisce o revoca un beneficio, concede o nega un titolo, irroga una sanzione – più complessa è la valutazione di ciò che può considerarsi effetti "significativi".

All'uopo, il considerando 71 del GDPR offre un primo orientamento, precisando che tali effetti sussistono quando l'automazione incide in modo rilevante sullo *status* giuridico o sulla condizione di vita del soggetto. Pur privo di valore vincolante, a ben vedere, questo chiarimento trova riscontro nelle FAQ della Commissione europea, secondo cui la disposizione si applica anche quando l'impatto del processo automatizzato non si traduce immediatamente in un effetto legale, ma altera in modo sostanziale le opportunità o le condizioni di vita dell'interessato.

Pensiamo a un algoritmo che classifica un contribuente come soggetto "ad alto rischio fiscale", il quale, pur non determinando direttamente un provvedimento, orienta la probabilità di controlli futuri, producendo effetti economici e psicologici non trascurabili. Analogamente, un sistema che attribuisce un *rating* negativo a un'impresa che intende partecipare a un appalto, non produce un effetto giuridico immediato, ma condiziona fortemente le decisioni dei commissari e, di riflesso, le possibilità di successo dell'operatore economico.

La Corte di giustizia dell'Unione europea, nella recente causa *Schufa* (C-634/21, 2023)<sup>70</sup>, ha confermato questa lettura estensiva, stabilendo che lo *score* creditizio elaborato da un'agenzia privata può configurare una decisione automatizzata ai sensi dell'art. 22 quando viene utilizzato in modo determinante da un terzo, come una banca, per concedere o negare un prestito.

Il principio affermato è di grande rilievo, poiché riconosce che anche valutazioni probabilistiche, se idonee a incidere concretamente sulle scelte che riguardano la persona, rientrano nella sfera di protezione della norma. L'impatto di questa pronuncia si riverbera direttamente sull'amministrazione predittiva, nella quale punteggi e classificazioni assumono un ruolo determinante nelle politiche di controllo, nell'erogazione di prestazioni e nella gestione delle risorse.

Nondimeno, pur alla luce del considerando 71, delle FAQ della Commissione europea e della giurisprudenza che ne hanno accolto la lettura

---

<sup>70</sup> Sul tema del diritto alla spiegazione, cfr. ora anche CGUE, C-203/22, *Dun & Bradstreet Austria GmbH* (2025), che, nel solco di *Schufa* (C-634/21), ha precisato che l'interessato, in caso di decisione automatizzata, può pretendere una spiegazione "significativa, comprensibile e accessibile" circa la logica utilizzata e che il bilanciamento con eventuali segreti commerciali deve essere rimesso all'autorità di controllo o al giudice competente

estensiva, la nozione di effetti “significativi” continua a difettare di una definizione normativa precisa, lasciando irrisolta una tensione di fondo tra certezza del diritto e adattabilità tecnologica.

Chiarito ciò, se la giurisprudenza e la prassi interpretativa hanno contribuito ad ampliare l’ambito di tutela dell’art. 22, la disposizione non rimane sgombra da ulteriori complicazioni, mostrando, ancora una volta, il suo volto ambiguo. Essa si colloca, infatti, nel punto di frizione tra la razionalità tecnico-decisionale propria dei sistemi automatizzati e la logica garantistica che ispira il diritto europeo alla protezione dei dati personali.

Tale tensione emerge in modo evidente nel par. 2 che attenua il divieto di decisioni unicamente automatizzate introducendo tre ipotesi di deroga<sup>71</sup>: quando la decisione è necessaria per la conclusione o l’esecuzione di un contratto, quando è autorizzata dal diritto dell’Unione o degli Stati membri e quando si fonda sul consenso esplicito dell’interessato.

In linea teorica, la clausola di deroga mira a bilanciare esigenze di efficienza e innovazione con la protezione dei diritti fondamentali. Tuttavia, nella pratica applicazione, la sua portata risulta ambigua e rischia di indebolire l’effettività della tutela.

Per comprendere meglio questa tensione, può essere utile soffermarsi sulla deroga più rilevante nel contesto amministrativo, ossia quella che consente agli Stati membri o al diritto dell’Unione di autorizzare decisioni automatizzate purché siano previste adeguate misure di salvaguardia. Da un lato, tale previsione permette di legittimare l’uso di algoritmi nei procedimenti fiscali, previdenziali o di polizia amministrativa; dall’altro, rischia di trasformarsi in una zona franca regolatoria, attraverso la quale si amplia la discrezionalità normativa senza che vi corrisponda un effettivo presidio garantistico.

Emergono, proprio in questo punto, alcune delle ricadute pratiche della difficile intersezione suggerite nei paragrafi precedenti. Le garanzie previste dall’art. 22, come il diritto all’intervento umano o la possibilità di contestare la decisione, restano infatti difficili da esercitare quando l’algoritmo è opaco e l’interessato non può accedere ai criteri che ne hanno determinato l’esito. Analogo discorso vale per la deroga fondata sul consenso esplicito, la cui effettiva libertà e consapevolezza sono state già messe in dubbio in ragione dell’asimmetria informativa e del condizionamento che caratterizzano l’interazione tra cittadino e amministrazione digitale.

È in questo contesto che si colloca l’altro nodo cruciale della disciplina, già tratteggiato nella riflessione sull’opacità algoritmica, ossia il cosiddetto “diritto alla spiegazione” (*right to explanation*) della decisione automatizzata, che rappresenta una logica derivazione delle problematiche emerse in materia di trasparenza, discrezionalità e accountability amministrativa.

Benché il GDPR non enunci espresamente un “diritto alla spiegazione”, diversi autori hanno sostenuto che la migliore lettura sistematica degli

---

<sup>71</sup> Per completezza, pur non costituendo il tema oggetto specifico della presente analisi, si richiama altresì l’ulteriore deroga prevista dall’art. 23 GDPR, che consente agli Stati membri di introdurre limitazioni per motivi di interesse pubblico rilevante. Tale disposizione è stata, ad esempio, applicata in Ungheria durante l’emergenza COVID-19, quando il diritto di opposizione alle decisioni automatizzate è stato temporaneamente sospeso per finalità emergenziali.

artt. 13, 15 e 22 ne sorregge l'esistenza in termini positivi, come pretesa a ricevere «*meaningful information about the logic involved*»<sup>72</sup> nei processi automatizzati, idonea a garantire l'esercizio effettivo dei diritti e un controllo sostanziale dell'esito.

Tuttavia, le opinioni, al riguardo, risultano divise. Altri autori, infatti, in un'ottica più restrittiva, negano l'esistenza di un diritto alla spiegazione in senso sostanziale, osservando che il GDPR riconosce solo obblighi informativi generali e garanzie procedurali; pertanto, «*after a decision has been made [...] data subjects are granted additional safeguards to obtain human intervention, express views, or contest a decision (Article 22(3)), but not to obtain an explanation of the decision reached*»<sup>73</sup>.

La questione ruota attorno al grado di comprensibilità della logica sottesa ai processi decisionali poiché informare non significa necessariamente spiegare e, in assenza di spiegazione, il diritto di difesa resta puramente formale.

In questo senso, secondo gli studiosi dell'*Oxford Internet Institute*<sup>74</sup>, auditi nel 2018 alla Camera dei Comuni britannica, vi sarebbe una lacuna nel GDPR: pur essendo ampiamente affermato in dottrina che l'art. 22 riconosca un diritto alla spiegazione delle decisioni algoritmiche, di fatto tale diritto non è esplicitamente previsto dalla disposizione vincolante.

Il prefato considerando 71, appunto, fa riferimento al diritto dell'interessato di ottenere spiegazioni sulla logica sottesa al trattamento automatizzato. Tuttavia, come già spiegato, il considerando «oltre ad essere una previsione giuridicamente non vincolante, per alcuni versi sembra limitarsi alla fumosa informazione circa l'esistenza di un procedimento informatizzato»<sup>75</sup>. Sotto altra angolazione, si potrebbe sempre osservare che l'art. 22, par. 3, impone al titolare del trattamento di adottare misure adeguate a salvaguardare i diritti e le libertà dell'interessato, garantendo almeno l'intervento umano, la possibilità di esprimere il proprio punto di vista e di contestare la decisione, ma non permette di ottenere una spiegazione dell'esito raggiunto. Tuttavia, come evidenziato dall'*Oxford Institute*, queste garanzie rischiano di essere di scarsa utilità se non accompagnate da un diritto effettivo a comprendere come la decisione sia stata presa.

Insomma, per garantire una reale effettività, e seguendo l'impostazione delle Linee guida dell'EDPB già citate, la trasparenza non implica l'accesso al codice sorgente, ma la possibilità di conoscere i criteri decisionali essenziali e il peso attribuito ai dati e ai fattori che hanno determinato l'esito.

La questione assume una rilevanza ulteriore in chiave costituzionale; infatti, se manca la spiegabilità, viene meno la possibilità stessa di verificare l'uguaglianza di trattamento, poiché eventuali discriminazioni o *bias* rimangono invisibili e, quindi, incontestabili.

Come già evidenziato nei paragrafi precedenti, l'AI Act si inserisce in que-

<sup>72</sup> A.D. Selbst-J. Powles, *Meaningful information and the right to explanation*, in *International Data Privacy Law*, 7-4, 2017, spec. 233 ss.

<sup>73</sup> S. Wachter-B. Mittelstadt-L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR*, in *International Data Privacy Law*, 7-2, 2017, spec. 79.

<sup>74</sup> Cfr. F. S. di S. Ippolito-M. Nicotra, *Diritto della blockchain, intelligenza artificiale e IoT*, Milano, 2018, spec. 270 ss.

<sup>75</sup> G. Alpa, *Diritto*, cit., 285.

sta prospettiva di rafforzamento, introducendo un approccio basato sul rischio che integra il GDPR e ne colma, almeno in parte, le lacune in tema di trasparenza, qualità dei dati e supervisione umana effettiva. (→ semplifico la sintassi per migliorare la fluidità) Tuttavia, più che rappresentare un punto di svolta, il nuovo Regolamento può essere letto come un completamento del sistema europeo di protezione dei dati, che trova nel GDPR il suo perno e nella combinazione dei due testi la chiave per bilanciare innovazione tecnologica e tutela dei diritti fondamentali.

Come si è già avuto modo di osservare, questo equilibrio si realizza soprattutto attraverso gli strumenti di accountability previsti dal GDPR, i quali assumono un rilievo determinante nel governo delle decisioni automatizzate. Essi traducono in pratica i principi di trasparenza e controllo umano, configurando un sistema di governance fondato su una chiara attribuzione di ruoli, obblighi e responsabilità tra i soggetti coinvolti nel trattamento dei dati personali.

Al centro di tale architettura si colloca la figura del titolare del trattamento, definito all'art. 4, par. 7, come la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che determina, da solo o congiuntamente ad altri, le finalità e i mezzi del trattamento. A questa figura compete la responsabilità primaria, anche quando le operazioni materiali siano delegate ad altri soggetti.

Ciò premesso, il principio di *accountability*, enunciato con chiarezza all'art. 5, par. 2, impone al titolare non solo di garantire la conformità alle disposizioni del GDPR, ma anche di poterne fornire prova concreta. Ne deriva un mutamento strutturale in cui non è più sufficiente "rispettare" il diritto, ma occorre dimostrare attivamente la conformità, attraverso strumenti organizzativi e procedurali che consentano la verifica ex post della correttezza delle operazioni di trattamento.

Tra gli strumenti di attuazione, la DPIA, prevista dall'art. 35, rappresenta il principale meccanismo di prevenzione dei rischi<sup>76</sup>. Essa impone al titolare un'attività anticipatoria di analisi e mitigazione, finalizzata a garantire trasparenza, proporzionalità e sicurezza.

In presenza di un rischio residuale elevato, il GDPR richiede, inoltre, una consultazione preventiva con l'autorità di controllo (art. 36), rafforzando l'impianto collaborativo e dinamico della protezione dei dati.

A fondamento della DPIA, e più in generale della responsabilità del titolare, si colloca il principio di liceità, correttezza e trasparenza (art. 5, par. 1, lett. a)), che impone una gestione chiara e leale dei dati fin dalle prime fasi del trattamento. Pertanto, ogni trattamento deve essere fondato su un presupposto giuridicamente legittimo, «conforme alle norme ad esso applicabili (liceità), nella prospettiva di assicurare un rapporto leale tra titolare e interessato (correttezza), avendo cura di raccogliere i dati per finalità

---

<sup>76</sup> Nel quadro dell'AI Act si individuano quattro categorie principali di rischio: inaccettabile, che implica il divieto assoluto di utilizzo dell'IA (cfr. art. 5 AI Act); alto, per il quale produttori e sviluppatori sono soggetti a rigorosi obblighi specifici che coprono l'intero ciclo di vita del sistema; basso; e, infine, una soglia specifica legata alla salvaguardia della trasparenza. Vd. M. C. Pollicino, *Gli effetti della "sommatoria" tra il GDPR e il nuovo Regolamento sulle intelligenze artificiali nell'ambito dell'attività amministrativa*, in *Rivista italiana di informatica e diritto*, 7, 2025, spec. 245 e 246.

determinate, esplicite e legittime (limitazione delle finalità)»<sup>77</sup>.

A supporto dell'*accountability*, l'art. 30 prevede l'obbligo di tenuta di un registro delle attività di trattamento, redatto e aggiornato a cura del titolare (e, ove previsto, del responsabile). Questo registro ha la funzione di rendere tracciabili tutte le operazioni effettuate, indicando le finalità, le categorie di dati e di interessati, i destinatari, i termini di conservazione e le misure di sicurezza adottate. Esso costituisce uno strumento imprescindibile per facilitare i controlli da parte delle autorità garanti e per dimostrare la conformità ai principi del Regolamento.

In tale cornice, anche il principio della protezione dei dati fin dalla progettazione (*privacy by design*) e per impostazione predefinita (*privacy by default*), previsto all'art. 25, assume un ruolo strategico. La *privacy by design* impone l'integrazione strutturale della protezione dei dati nei processi aziendali e tecnologici, sin dalle fasi di progettazione dei sistemi e dei servizi. La *privacy by default* richiede invece che, per impostazione predefinita, siano trattati solo i dati strettamente necessari rispetto alle finalità perseguite. Entrambi i principi impongono una logica di prevenzione e contenimento del rischio, che risulta cruciale soprattutto nei trattamenti automatizzati, come quelli che impiegano algoritmi di intelligenza artificiale.

Infine, la complessità delle operazioni di trattamento rende necessaria la presenza del Responsabile della Protezione dei Dati (DPO), obbligatorio in determinati casi previsti dall'art. 37. Il DPO assume funzioni di sorveglianza, consulenza e interfaccia con le autorità di controllo, garantendo il rispetto delle norme e fungendo da presidio interno di legalità. La sua presenza è particolarmente rilevante nei contesti di trattamento ad alto rischio, dove la valutazione e la gestione del rischio devono essere svolte in modo indipendente, continuativo e professionale.

Per riepilogare ed avvicinarci alle conclusioni, la DPIA si configura come uno degli strumenti più rappresentativi del nuovo approccio proattivo e dinamico alla protezione dei dati personali: se da una parte, evidenzia l'effettività del GDPR di giocare sul terreno della valutazione *ex ante* dei rischi, rappresentando il fulcro di una cultura della protezione fondata sulla responsabilizzazione concreta dei soggetti coinvolti; dall'altra parte, essa esprime in modo paradigmatico il passaggio da una logica meramente autorizzativa a una responsabilità sostanziale, orientata alla prevenzione, alla trasparenza e alla rendicontazione.

Per queste ragioni, sempre più si discute della necessità di integrare la DPIA con valutazioni più ampie, orientate ai diritti fondamentali (FRIA), in grado di considerare anche gli impatti sociali e non strettamente *privacy-related* dei sistemi predittivi.

Le tecnologie di protezione della privacy (PETs), come la crittografia, la pseudonimizzazione, l'apprendimento federato e la *differential privacy*, offrono ulteriori strumenti per ridurre i rischi.

A questi si aggiungono gli *audit* indipendenti, che consentono di verificare nel tempo l'assenza di bias e la conformità del modello ai principi normativi, e i meccanismi di supervisione umana effettiva, che devono essere

---

<sup>77</sup> S. Franca, *Il trattamento dei dati nelle sperimentazioni di intelligenza artificiale riguardanti le Pubbliche Amministrazioni*, in *Intelligenza artificiale e diritto: una rivoluzione?*, Quaderni ASTRID, vol. 2, Bologna, 2022, spec. 167.

pensati non come mero presidio formale, ma come funzione organizzativa dotata di poteri reali di controllo e intervento.

In definitiva, l'art. 22 GDPR rappresenta una clausola di protezione essenziale, ma non autosufficiente. La sua efficacia dipende dall'interpretazione estensiva dei concetti chiave, dalla sostanza delle garanzie applicate nelle deroghe e dall'effettività del diritto alla spiegazione. La giurisprudenza recente, come il caso *Schufa*, indica un orientamento verso la massimizzazione della tutela, ma resta la necessità di strumenti ulteriori per rendere effettiva la protezione contro i rischi della predizione amministrativa.

L'AI Act e gli strumenti di accountability già previsti dal GDPR si collocano in questa prospettiva di rafforzamento, ma il loro successo dipenderà dalla capacità degli Stati membri di tradurli in pratiche organizzative solide e trasparenti.

## 5. Conclusioni

L'amministrazione predittiva non è illegittima per natura, ma lo diventa quando l'*output* sostituisce il giudizio, l'opacità erode la motivazione e l'imputabilità si dissolve in automatismi non verificabili. In questa prospettiva, una lettura sostanziale dell'art. 22 del GDPR, coordinata con i doveri di supervisione umana effettiva introdotti dall'AI Act, conduce a un esito nitido: la spiegabilità funzionale, intesa come condizione giuridica di legalità, diventa il punto di equilibrio tra potere e controllo, perché consente di comprendere i fattori determinanti della decisione e il loro peso "a misura di destinatario". Da ciò discende che il consenso non può fungere da base ordinaria nei procedimenti pubblici, essendo strutturalmente fragile in presenza di asimmetrie informative.

La garanzia deve, invece, collocarsi nel cuore dell'organizzazione amministrativa, attraverso un sistema di responsabilità progettuale e continuativa che renda verificabile ogni fase del ciclo algoritmico. Valutazioni d'impatto più ricche (DPIA e, ove opportuno, FRIA), *fairness testing*, *audit* indipendenti e tracciamento delle revisioni e delle eventuali disattese dell'*output* rappresentano, in questa logica, strumenti di imputabilità sostanziale prima ancora che di conformità formale.

Inoltre, poiché gli algoritmi apprendono correlazioni e tendono a cristallizzare disuguaglianze sedimentate nei dati, la spiegabilità si salda naturalmente con il principio di eguaglianza sostanziale, che torna a svolgere la funzione di bussola delle scelte tecniche. L'eguaglianza, in tal senso, non si limita a correggere gli esiti della predizione, ma ne orienta la costruzione stessa, imponendo che le metriche di equità siano sottoposte a un controllo di proporzionalità e di ragionevolezza. Ne deriva una responsabilità amministrativa di tipo sistemico, che non si esaurisce nell'atto singolo, ma percorre l'intero ciclo di vita del modello, dalla progettazione all'uso operativo e alla sua revisione.

In questo quadro, la predizione non sostituisce la decisione, né l'efficienza algoritmica può surrogare la giustificazione pubblica: la tecnologia rimane strumento e l'esercizio del potere conserva il suo carattere umano, comprensibile e responsabile. Inoltre, l'eguaglianza, nell'era dell'intelligenza ar-

tificiale, non è un vincolo esterno ma la condizione che impedisce alla previsione di diventare destino; la legalità, per non ridursi a formula rituale, deve farsi sostanziale nella convergenza tra protezione dei dati, spiegabilità e imputabilità. Solo così l'amministrazione predittiva potrà dirsi, insieme, efficiente e giusta.

### **Abstract**

Il contributo analizza le implicazioni giuridiche dell'impiego di sistemi predittivi nella pubblica amministrazione nel campo della protezione dei dati personali, mettendo in luce le tensioni che essi generano rispetto ai principi di legalità, trasparenza, responsabilità ed eguaglianza. Muovendo da una lettura sostanziale dell'art. 22 GDPR, integrata con gli obblighi di supervisione umana introdotti dall'AI Act, il saggio vuole dimostrare come l'automazione decisionale rimodelli il potere amministrativo e metta in discussione la sua intelligibilità e imputabilità. L'analisi, quindi, propone una ricostruzione dogmatica fondata sull'idea che la responsabilità organizzativa e il controllo umano costituiscano condizioni essenziali di una legalità algoritmica dell'azione amministrativa, nella quale la protezione dei dati e l'eguaglianza sostanziale convergono in un'unica architettura di garanzie.

The paper analyses the legal implications of the use of predictive systems within public administration in the field of personal data protection, highlighting the tensions they generate with regard to the principles of legality, transparency, accountability, and substantive equality. Building on a substantive reading of Article 22 of the GDPR, integrated with the human oversight obligations introduced by the AI Act, the study aims to demonstrate how automated decision-making reshapes administrative power and calls into question its intelligibility and imputability. The analysis therefore proposes a dogmatic reconstruction based on the idea that organisational accountability and human control constitute essential conditions for the algorithmic legality of administrative action, in which data protection and substantive equality converge within a single architecture of safeguards.

### **Keywords**

amministrazione algoritmica – art. 22 GDPR – AI Act – spiegabilità funzionale – eguaglianza sostanziale