

## Intelligenza artificiale e pubbliche amministrazioni: affrontare i numerosi rischi per trarne tutti i vantaggi

Nicoletta Rangone \*

ARTIFICIAL INTELLIGENCE AND PUBLIC ADMINISTRATIONS: ADDRESSING THE MANY RISKS TO GAIN ALL THE BENEFITS

ABSTRACT: The paper aims at analysing the most relevant examples of the use of artificial intelligence within the Italian public administrations, in order to highlight related benefits and risks. It focuses on the internal organization, rulemaking and enforcement decisions. There appears to be a need for a general regulatory framework, as well as for specific procedural disciplines in each administration in order to allow real accessibility, comprehensibility and non-discrimination for the stakeholders. At the same time, these disciplines should introduce mechanisms and evaluation adequate to verify the appropriateness of the use of artificial intelligence instead of human intelligence, with a view to the effectiveness of organization and administrative action.

KEYWORDS: artificial intelligence; administrative decision; regulation; transparency; cognitive bias

ABSTRACT: Lo scritto parte dall'analisi delle più rilevanti esperienze di uso dell'intelligenza artificiale nelle pubbliche amministrazioni italiane, che attengono all'assetto organizzativo, all'adozione di decisioni generali e all'attuazione amministrativa, per mettere in luce vantaggi e rischi. Ne emerge l'esigenza di un inquadramento normativo generale e di discipline procedurali delle singole amministrazioni, tali da assicurare reale accessibilità, comprensibilità e non discriminazione, oltre a meccanismi di controllo che consentano di verificare l'adeguatezza del ricorso all'intelligenza artificiale in luogo dell'intelligenza umana in un'ottica di effettività dell'organizzazione e dell'azione amministrativa.

---

\* Professoressa di diritto amministrativo, Università Lumsa. Mail: [n.rangone@lumsa.it](mailto:n.rangone@lumsa.it). Una prima versione del contributo è destinata alla pubblicazione in A. LALLI (a cura di), *Le amministrazioni pubbliche nell'era digitale*, Giappichelli in corso di pubblicazione. Le considerazioni svolte sono frutto della ricerca avviata nell'ambito del progetto PRIN 2017 "governance of/through Big Data: challenges for european law" e proseguita nel gruppo di ricerca ASTRID «Amministrazione e Intelligenza Artificiale» coordinato con B. Marchetti ed E. Chiti. L'autore è titolare della cattedra Jean Monnet on EU approach to Better regulation; il sostegno della Commissione europea alla produzione di questa pubblicazione non costituisce un'approvazione del contenuto, che riflette esclusivamente il punto di vista dell'autore, e la Commissione non può essere ritenuta responsabile per l'uso che può essere fatto delle informazioni ivi contenute. Contributo sottoposto a doppio referaggio anonimo.

PAROLE CHIAVE: intelligenza artificiale; decisioni amministrative; regolazioni; trasparenza; bias comportamentali

SOMMARIO: 1. Introduzione – 2. Intelligenza artificiale per ottimizzare prestazioni e organizzazione delle amministrazioni pubbliche – 3. Intelligenza artificiale per il *rule-making* – 4. Intelligenza artificiale per l'attuazione amministrativa – 4.1. Trasparenza effettiva e spiegabilità: obiettivi raggiungibili? – 4.2 Non esclusività della decisione algoritmica: criticità – 5. Esigenza di un quadro normativo minimo, da specificare a livello di singola amministrazione.

## 1. Introduzione

Diverse e sempre più diffuse sono le applicazioni di intelligenza artificiale<sup>1</sup> nelle pubbliche amministrazioni italiane, a supporto di servizi al pubblico, di esigenze organizzative e conoscitive. Le nuove tecnologie sono inoltre utilizzate nei processi decisionali, o nella relativa fase prodromica, che si tratti dell'adozione di *policies*, di regole o di decisioni amministrative.

Questi usi di nuove tecnologie sono all'origine di numerosi vantaggi per i pubblici poteri, che vanno dalla possibilità di comprimere i tempi per decidere e di razionalizzare l'impegno di risorse umane, alla elaborazione di indicazioni anche prospettiche, alla possibilità di evitare errori umani e di limitare le occasioni di corruzione<sup>2</sup>. L'intelligenza artificiale costituisce dunque uno strumento per l'effettività dell'organizzazione e dell'azione amministrativa, in ultima analisi di un diritto amministrativo che sia osservato, attuato e foriero di risultati coerenti con gli obiettivi per cui è stato previsto<sup>3</sup>. Un importante stimolo all'uso dell'intelligenza artificiale nell'organizzazione e azione amministrativa è svolto dagli

<sup>1</sup> In tal contesto, si può richiamare la definizione di intelligenza artificiale offerta dal Consiglio di Stato, per cui il *machine learning* «crea un sistema che non si limita solo ad applicare le regole software e i parametri preimpostati (come fa invece l'algoritmo "tradizionale") ma, al contrario, elabora costantemente nuovi criteri di inferenza tra dati e assume decisioni efficienti sulla base di tali elaborazioni, secondo un processo di apprendimento automatico» (punto 9.1, Consiglio di Stato, sez. III, 25 novembre 2021, n. 7891). Quanto alla definizione di intelligenza artificiale elaborata in seno alla proposta di Regolamento europeo sull'intelligenza artificiale (considerando 6 e art. 3), si rinvia a G. AVANZINI, *Intelligenza artificiale, machine learning e istruttoria procedimentale: vantaggi, limiti ed esigenza di una specifica data governance*, in F. DONATI, A. PAJNO, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, Bologna, Il, 2022, cap. 2, 47-48.

<sup>2</sup> Nella nota vicenda della mobilità docenti, il Consiglio di Stato (sez. VI, 4 febbraio 2020, n. 881) ha evidenziato che «l'utilizzo di una procedura informatica che conduca direttamente alla decisione finale non deve essere stigmatizzata, ma anzi, in linea di massima, incoraggiata: essa comporta infatti numerosi vantaggi quali, ad esempio, la notevole riduzione della tempistica procedimentale per operazioni meramente ripetitive e prive di discrezionalità, l'esclusione di interferenze dovute a negligenza (o peggio dolo) del funzionario (essere umano) e la conseguente maggior garanzia di imparzialità della decisione automatizzata».

<sup>3</sup> G. CORSO, M. DE BENEDETTO, N. RANGONE, *Il diritto amministrativo effettivo. Una introduzione*, Bologna, 2022, 15. Il capitolo 3 di tale testo sviluppa una prima articolazione delle argomentazioni svolte in questo lavoro.



incentivi europei di tipo economico<sup>4</sup> (da ultimo il *Recovery and Resilience Facility*)<sup>5</sup> e non economici che derivano dal confronto delle pratiche in sede internazionale<sup>6</sup> ed europea<sup>7</sup>.

Gli indubbi vantaggi connessi all'uso delle nuove tecnologie non possono adombrare però i rischi, che vanno dall'opacità, all'errore, alla discriminazione<sup>8</sup>. Ed invero, se l'intelligenza artificiale nel processo decisionale consente di ridurre il «rumore» (*noise*), vale a dire la «variabilità» ingiustificata nei giudizi umani su situazioni identiche<sup>9</sup>, può essa stessa essere il veicolo attraverso il quale sono perpetrate «le storture e le imperfezioni che caratterizzano tipicamente i processi cognitivi e le scelte compiute dagli esseri umani»<sup>10</sup> (*garbage in, garbage out*)<sup>11</sup>. Distorsioni possono anche essere legate alla scarsa qualità dei dati che alimentano l'intelligenza artificiale, ad esempio perché raccolti da soggetti in conflitto di interessi, basati su presunzioni non validate dal diretto interessato o esclusivamente su dati storici (che da un lato lasciano «sotto il radar» i nuovi entranti e nuovi rischi, dall'altro sono usati sulla base dell'assunto che il comportamento osservato non cambi in futuro<sup>12</sup>), raccolti per scopi diversi oppure tratti da «tracce» lasciate in rete dalle persone<sup>13</sup>. Ed ancora, un modello potrebbe risultare impreciso quando utilizzato nel mondo reale<sup>14</sup>.

Vi sono poi problematiche specifiche, che attengono ai diversi momenti ed obiettivi per i quali i pubblici poteri fanno uso dell'intelligenza artificiale, su cui si concentra il presente contributo, che invece non affronta il tema delle competenze. Il paragrafo 2 è dedicato all'uso dell'intelligenza artificiale per il miglioramento delle prestazioni al pubblico e l'organizzazione interna, il paragrafo 3 alle nuove tecnologie nei procedimenti di regolazione e normativi, il paragrafo 4 all'attuazione amministrativa. Le

<sup>4</sup> Ad esempio, la *detection* algoritmica del linguaggio d'odio da parte dell'Autorità per le garanzie nelle comunicazioni ha ricevuto impulso anche da un progetto finanziato dalla Commissione europea (*Innovative Monitoring Systems and Prevention Policies of Online Hate Speech-IMSYP*) che ha portato l'autorità a lavorare nell'ambito di un consorzio di ricerca internazionale.

<sup>5</sup> Al rispetto e all'implementazione del principio, tra gli altri, del *digital by default* sono vincolate le riforme supportate dal *Recovery and Resilience Facility*.

<sup>6</sup> Ad esempio, il premio UNESCO al sistema INPS di smistamento delle PEC. Ed ancora, gli scrutini periodici di *mutual evaluation* realizzati a livello OCSE nel Gruppo intergovernativo d'Azione Finanziaria Internazionale-GAFI sono l'occasione per lo scambio di buone pratiche anche con riferimento all'uso dell'intelligenza artificiale nella vigilanza di Banca d'Italia. Questi esempi sono riportati da E. CHITI, B. MARCHETTI, N. RANGONE, *L'impiego di sistemi di intelligenza artificiale nelle pubbliche amministrazioni italiane: prove generali*, in F. DONATI, A. PAJNO, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., II, cap. 1, e in questa rivista.

<sup>7</sup> Ad esempio, in ambito ESMA.

<sup>8</sup> In tema, si veda F. COSTANTINO, *Rischi o opportunità del ricorso delle amministrazioni alle predizioni dei big data*, in *Diritto pubblico*, 1, 2019, 43 ss.

<sup>9</sup> «Whether the patent office grants or rejects a patent is significantly related to the happenstance of which examiner is assigned the application. This variability is obviously troublesome from the standpoint of equity» (D. KAHANEMAN, O. SIBONY, C.R. SUNSTEIN, *Noise. A Flaw in human judgment*, New York, 2021, 6-7).

<sup>10</sup> Consiglio di Stato, VI, 13 dicembre 2019 n. 8472.

<sup>11</sup> S.G. MAYSON, *Bias In, Bias Out*, in *The Yale Law Journal*, 128, 2019, 2218 ss.

<sup>12</sup> Definita «illusion of validity» su cui *infra*.

<sup>13</sup> T.B. GILLIS, *The input fallacy*, in *Minnesota Law Review*, 106, 2022, 1175 ss.

<sup>14</sup> «First, a model can simply fail to fit any data — training or test — well. In such a scenario, even if the training and test data were perfectly representative of real-world data, the model would be inaccurate when deployed. Second, an algorithm can fit its training and, perhaps, test data well, but fail to generalize and perform equally well in real-world data» (D. LEHR, P. OHM, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, in *University of California*, 51, 2017, 711).



conclusioni, di cui al paragrafo 5, sollevano l'esigenza di un inquadramento normativo per gli usi dell'intelligenza artificiale basato sulla trasparenza (in termini di effettiva conoscibilità), come regola minima e comune a tutte le applicazioni, ed esteso ai presidi della motivazione (come effettiva spiegabilità) e verificabilità dei sistemi per gli usi da parte delle pubbliche amministrazioni nei processi decisionali pubblici.

## 2. Intelligenza artificiale per ottimizzare prestazioni e organizzazione interna delle amministrazioni pubbliche

L'intelligenza artificiale viene utilizzata per ottimizzare le prestazioni al pubblico e l'organizzazione interna.

Quanto ai servizi al pubblico, alcune applicazioni predisposte o in via di sperimentazione sono volte a facilitare il rapporto con i cittadini, rendendo più efficaci le funzioni di prestazione.

Pur restando fondamentale la possibilità di interfacciarsi con una persona fisica<sup>15</sup>, di grande utilità appare il miglioramento del supporto all'utenza che può essere offerto da chatbot "intelligenti" in grado di comprendere le richieste di cittadini diversi per preparazione e provenienza geografica (in uso ad esempio da parte di INPS e Agenzia delle entrate)<sup>16</sup>. Nella medesima ottica di facilitare il dialogo con l'amministrazione, muovono le applicazioni che attengono al miglioramento della funzionalità di motori di ricerca (anch'essi "intelligenti") di siti istituzionali, riducendo le barriere all'accesso e indirizzando l'utente verso pagine di effettiva utilità (come nella sperimentazione INPS). I sistemi che apprendono dalle ricerche fatte dai cittadini consentono, ad esempio, di selezionare le parole che consentono di raggiungere un determinato servizio all'interno del sito istituzionale (come può essere l'indicazione di "mensa scolastica" in luogo di "refezione"). Questi usi dell'intelligenza artificiale (come anche quelli che consentono la traduzione automatica da e verso varie lingue straniere o di filtrare messaggi di posta elettronica che costituiscono spam) potenziano i risultati ottenibili attraverso sistemi deduttivi alimentati attraverso un numero limitato di dati e presentano bassissimi livelli di rischio (in termini di gravità e probabilità) per la riservatezza dei dati raccolti dagli utenti e altri diritti fondamentali.

Una seconda tipologia di servizi al pubblico attiene alle numerose applicazioni che caratterizzano le *smart cities*<sup>17</sup>, come la resa di informazioni raccolte attraverso l'internet delle cose e fornite agli utenti tramite app o siti istituzionali in ordine, ad esempio, alla disponibilità di parcheggi, alla densità di frequentazione di una determinata area di una città, a pericoli presenti sul manto stradale (in

<sup>15</sup> C. COGLIANESE, *Administrative law in the automated State*, in *American Academy of Arts and Sciences*, 2021, 204 ss.; S. RANCHORDÁS, *Empathy in the digital administrative law*, in *Duke Law Journal*, 71, 1341 ss.; J. PONCE SOLÉ, *Inteligencia artificial, Detecho admiistrativo y reserve de humanidad: algoritmos y procedimiento administrative debito tecnologico*, in *Revista General de Derecho Administrativo*, 50, 2019.

<sup>16</sup> G. BUONO, P. BONANNI, G. DEL MONDO, A. CIRIELLO, *Intelligenza artificiale e amministrazioni centrali*, in *Biolaw Journal*, 1, 2022, 261 ss.

<sup>17</sup> E. CHITI, B. MARCHETTI, N. RANGONE (a cura di), *L'uso dell'intelligenza artificiale nel sistema amministrativo italiano*, rapporto 3/2022, *SMART cities e intelligenza artificiale*, in *BioLaw Journal*, 1, 2022, 251 ss. Sull'uso dell'internet delle cose e l'intelligenza artificiale per offrire informazioni più puntuali, quando non personalizzate, agli utenti e adeguare l'organizzazione dell'offerta, si veda OCSE, *Shaping the future of regulators. The impact of emerging technologies on economic regulators*, 2020.

sperimentazione in alcuni comuni, come Padova, Trento e Venezia)<sup>18</sup>. Si tratta di funzionalità che potrebbero presentare rischi, soprattutto in termini di tutela privacy, là dove i dati siano raccolti attraverso telecamere e sensori acustici<sup>19</sup>. È dunque importante, in queste ipotesi, non solo assicurare la trasparenza degli usi, ma anche – ove possibile – evitare a monte questi rischi anonimizzando i dati proprio attraverso l'intelligenza artificiale e registrando solo il metadato (come nelle esperienze di Padova e Venezia), oppure creando un *blind learning environment*, che consente di evitare la condivisione di dati “in chiaro” all'interno di un medesimo soggetto pubblico (come nel caso di Banca d'Italia)<sup>20</sup>.

Altri usi dell'intelligenza artificiale attengono all'organizzazione interna delle pubbliche amministrazioni e consentono di ottimizzare la distribuzione del lavoro tra uffici o la classificazione automatica della posta in entrata (come sperimentato dall'INPS)<sup>21</sup>. Si tratta di applicazioni che non sembrano presentare particolari rischi per i dipendenti o i soggetti terzi, se non in termini di una eventuale inefficiente organizzazione del lavoro o ritardati servizi all'utenza.

Diverso è il caso della razionalizzazione di attività che comportino la raccolta di informazioni che possono risultare prodromiche all'avvio eventuale di un procedimento di controllo, regolatorio o decisionale. Si pensi alla categorizzazione di esposti e denunce effettuata da Banca d'Italia<sup>22</sup> e alla vigilanza bancaria per la rilevazione di anomalie indicative di possibili azioni di riciclaggio<sup>23</sup>, alla classificazione per grado di rischio di documenti oggetto di vigilanza (come nel caso documenti sintetici – KID – che illustrano le caratteristiche dei prodotti finanziari PRIIPs rivolti agli investitori al dettaglio oggetto di

<sup>18</sup> Altre applicazioni possibili, in corso di realizzazione in altri ordinamenti giuridici, attengono alla robotica e consentono, ad esempio, la consegna di medicinali attraverso droni o la corrispondenza attraverso mezzi a guida autonoma (D.F. ENGSTROM, D.E. HO, C.M. SHARKEY, M.F. CUELLAR, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, Report submitted to the Administrative Conference of the United States, 65-69. Per un commento al report, si veda L. Parona, “*Government by algorithm*” un contributo allo studio del ricorso all'intelligenza artificiale nell'esercizio di funzioni amministrative, in *Giornale di diritto amministrativo*, 1, 2021, 10 ss.).

<sup>19</sup> Ed invero, la proposta di regolamento europeo sull'intelligenza artificiale elenca come ad alto rischio «i sistemi di IA destinati a essere utilizzati per l'identificazione biometrica remota “in tempo reale” e “a posteriori” delle persone fisiche» (allegato III).

<sup>20</sup> E. CHITI, B. MARCHETTI, N. RANGONE, *L'impiego di sistemi di intelligenza artificiale nel sistema amministrativo italiano: prove generali*, cit.

<sup>21</sup> G. BUONO, P. BONANNI, G. DEL MONDO, A. CIRIELLO, *Intelligenza artificiale e amministrazioni centrali*, cit.

<sup>22</sup> Regolamento sul trattamento dei dati personali nella gestione degli esposti del 2022, pubblicato in gazzetta ufficiale e sul sito di Banca d'Italia. In base a tale regolamento, seppure la segnalazione non porti all'avvio un procedimento amministrativo, colui che presenta l'esposto e gli intermediari potenzialmente interessati devono sapere che le informazioni acquisite vengono trattate attraverso *machine learning*. Devono poi essere predisposte tecniche per garantire il rispetto della riservatezza dei dati personali, la sorveglianza umana del monitoraggio e dell'aggiornamento delle tecniche di *machine learning* (anche per assicurarne la spiegabilità), l'assenza di uso per decisioni automatizzate, profilazione o predizione di comportamenti. Questi gli aspetti particolarmente apprezzati dal Garante della privacy, nel parere del 24 febbraio 2022 alla Banca d'Italia sullo schema di regolamento concernente il trattamento dei dati personali effettuato nell'ambito della gestione degli esposti.

<sup>23</sup> E. CHITI, B. MARCHETTI, N. RANGONE (a cura di), *L'uso dell'intelligenza artificiale nel sistema amministrativo italiano*, rapporto 2/2021, *L'impiego dell'intelligenza artificiale nell'attività di Banca d'Italia*, in *BioLaw Journal*, 4, 2021, 229 ss.





controllo da parte della Consob)<sup>24</sup>, al monitoraggio dell'offerta abusiva *on line* di attività finanziarie riservate effettuato da Consob, o dell'uso del *data mining* e *machine learning* da parte di Agenzia delle entrate e INPS per rilevare i fenomeni fraudolenti nella lotta all'evasione contributiva. Nella valutazione di esposti e denunce (di cui si dirà nel paragrafo successivo) il rischio consiste in un possibile impoverimento del bagaglio informativo del decisore pubblico se non adeguatamente valutate/classificate, oltre a violazioni della privacy.

L'uso dell'intelligenza artificiale per la programmazione dei controlli informati al rischio (come nelle sperimentazioni supportate dall'OCSE e che coinvolgono la provincia autonoma di Trento e le regioni Lombardia e Campania)<sup>25</sup>, che consente controlli più mirati anche in funzione prospettica, la eventuale limitata qualità dei dati utilizzati (come nel caso dell'uso di "data from the wild")<sup>26</sup>, unitamente alla mancanza di trasparenza sul loro impiego, potrebbe non consentire ai soggetti interessati di contestare, ad esempio, l'erroneità del dato utilizzato. Inoltre, seppure un controllo non comporti necessariamente una conseguenza giuridica (in termini ad esempio di sanzione), esso rappresenta di per sé stesso un costo quando effettuato in loco; peraltro, la rilevazione di una anomalia anche non grave costituirà un precedente preso in considerazione nella mappatura del rischio che caratterizza un determinato individuo o impresa, comportando l'esposizione a nuovi controlli futuri. È dunque importante che i dati siano verificabili e trasparenti e che gli esiti dei controlli "algoritmici" siano costantemente monitorati in modo da "allenare" ed ricalibrare l'algoritmo ove necessario. Il ricorso a controlli "algoritmici" dovrebbe essere reso noto al pubblico (trasparente) nelle forme che verranno evidenziate nel paragrafo 5.

L'intelligenza artificiale può, poi, supportare l'organizzazione di controlli in un'ottica di ordine pubblico, come nel caso di applicazioni utilizzate da alcune questure italiane che processano dati raccolti da telecamere dislocate sul territorio e dati in possesso della Polizia, del Ministero dell'interno o altri soggetti pubblici (in alcuni contesto anche social media) per aumentare l'effettività dei controlli, attraverso una loro programmazione in base al rischio. Oltre ai pericoli menzionati, le applicazioni di polizia predittiva, seppure nel nostro ordinamento giuridico limitate a tracciare «mappe» del rischio<sup>27</sup> e non «persone a rischio»<sup>28</sup>, potrebbero riprodurre pregiudizi tipici della valutazione umana rispetto a certe

<sup>24</sup> E. CHITI, B. MARCHETTI, N. RANGONE (a cura di), *L'uso dell'intelligenza artificiale nel sistema amministrativo italiano*, rapporto 1/21, *L'impiego dell'intelligenza artificiale nell'attività di CONSOB, Agcom e ARERA*, in *Biolaw*, 4, 2021, 215.

<sup>25</sup> F. BLANC, M. BENEDETTI, C. BERTONE, *L'informatica e il machine learning al servizio della semplificazione dei controlli sulle imprese: un equilibrio ancora da definire*, in *Studi parlamentari e di politica costituzionale*, 209, 2021, 121 ss.

<sup>26</sup> Si tratta dell'uso di dati raccolti a fini diversi o prodotti in esito ad altre attività che vanno ad alimentare sistemi di intelligenza artificiale (che da questi apprendono) senza che vi sia stato un riscontro esplicito e consapevole da parte del proprietario del dato o un monitoraggio da parte di terzi sulla qualità dello stesso (N. CRISTIANINI, *Shortcuts to Artificial Intelligence*, in M. PELILLO, T. SCANTAMBURLO (a cura di), *Machines We Trust: Perspectives on Dependable AI*, Cambridge, Massachusetts, London, England, 2021, 15-17).

<sup>27</sup> Sulle esperienze di alcune questure italiane che utilizzano sistemi di polizia predittiva KeyCrime, X-Law, E-Security si rinvia a M.B. ARMIENTO, *La polizia predittiva come strumento di attuazione amministrativa delle regole*, in *Diritto amministrativo*, 4, 2020, 990-991 e, dello stesso autore, *Nuove tecnologie e "nuova" sicurezza delle città*, in *Studi parlamentari e di politica costituzionale*, 209, 2021, 95 ss.

<sup>28</sup> I sistemi che interessano le persone nell'ambito delle attività di contrasto di reati sono considerati ad alto rischio dalla proposta di regolamento europeo, allegato III.

zone e dunque persone che vi abitano<sup>29</sup> e, là dove diano un peso prevalente ai dati storici, possono portare a concentrare i controlli su certi tipi di crimini o aree perdendone di vista altre o nuovi rischi, se non a stigmatizzare determinate persone. Questi sistemi, pur non essendo considerati a livello europeo (forse paradossalmente) tra quelli ad alto rischio, devono comunque essere informati al principio della trasparenza in base alla proposta di regolamento sull'intelligenza artificiale. A ciò si aggiunga il monitoraggio dei risultati e la verifica del buon funzionamento dell'algoritmo.

### 3. Intelligenza artificiale per il *rule-making*

Nei procedimenti di regolazione e normativi, le nuove tecnologie sono utili per riorganizzare e analizzare i commenti raccolti nell'ambito di consultazioni con un elevato numero di partecipanti<sup>30</sup>, per supportare la raccolta e l'elaborazione di dati di varia provenienza (come denunce e reclami) che possono far emergere l'esigenza di un intervento di regolazione<sup>31</sup>, per svolgere attività automatizzata o semi-automatizzata di *drafting*.

Nel nostro ordinamento non sembrano esservi esperienze di uso del *machine learning* per la lettura dei risultati di consultazioni, fatta eccezione per due politiche del 2014, «La buona scuola»<sup>32</sup> e «Rivoluzione@governo»<sup>33</sup>. I commenti ricevuti nella seconda consultazione sono stati il risultato di diverse campagne di mobilitazione promosse da soggetti organizzati per l'invio di commenti identici o sostanzialmente tali, di seguito definiti commenti di massa<sup>34</sup>. Nessuna applicazione in questo ambito si rileva invece, al momento, nelle autorità indipendenti italiane, che pure vantano le esperienze più avanzate in termini di consultazioni, presumibilmente per il numero contenuto di commenti normalmente ricevuto.

Quali considerazioni si possono svolgere sull'uso dell'intelligenza artificiale nelle consultazioni? A fronte di un prezioso supporto offerto ai decisori, l'uso in questa fase delle tecnologie più avanzate presenta rischi specifici, di cui occorre avere consapevolezza affinché possano essere adeguatamente affrontati dagli sviluppatori e dai decisori pubblici. Nella trattazione dei commenti di massa, l'intelligenza artificiale consente di identificare<sup>35</sup> e riorganizzare le osservazioni ricevute, generando un

<sup>29</sup> P.J. BRANTINGHAM, M. VALASIK, G.O. MOHLER, *Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial*, in *Statistics and Public Policy*, 5, 1, 2017, 1 ss.

<sup>30</sup> Ad esempio, nell'ambito delle consultazioni realizzate per l'analisi preventiva di impatto della proposta di direttiva europea su *Energy performance of buildings (recast)* «the results of the feedback were analysed using Atlas.ti (text processing software)» (SWD(2021) 453 final, 142).

<sup>31</sup> M.A. LIVERMORE, V. EIDELMAN, B. GROM, *Computational assisted regulatory participation*, in *Notre Dame Law Review*, 2018, 93, 3, 977 ss.

<sup>32</sup> Nella consultazione, svoltasi dal 15 settembre al 15 novembre 2014, le 6.470.000 risposte strutturate ricevute sono state analizzate, per l'estrazione di concetti chiave, attraverso la linguistica computazionale con il supporto della Fondazione Bruno Kessler.

<sup>33</sup> Le 39.343 mail pervenute nell'ambito della consultazione che si è svolta dal 30 aprile – 30 maggio 2014 sono state analizzate attraverso strumenti di *text mining* per classificare i messaggi ricevuti secondo il grado di pertinenza con i 44 punti della riforma, con il supporto del Dipartimento di metodi e modelli per l'economia, il territorio e la finanza dell'Università Sapienza.

<sup>34</sup> «As a rule of thumb, the minimum threshold should be 10 or more identical responses (across all the closed questions) to count as a campaign» (European Commission, *Better Regulation Toolbox*, 2021, 473).

<sup>35</sup> Sui software dedicati si veda European Commission, *Better Regulation Toolbox*, 2021, 473.



considerevole risparmio di tempo all'amministrazione che conduce la consultazione<sup>36</sup>. Questo sistema di razionalizzazione dell'analisi rischierebbe però di alterare il bagaglio informativo su cui si basa il processo decisionale se supportasse l'automatica esclusione dei commenti di massa in quanto tali o la loro considerazione come un unico commento<sup>37</sup>. Lo stesso è da dirsi per l'uso delle nuove tecnologie per semplificare l'analisi di un'importante mole di dati frutto di segnalazioni, denunce o reclami (sperimentata da Consob e Banca d'Italia)<sup>38</sup>, là dove ricevessero un'attenzione minore perché, ad esempio, contenenti errori ortografici o un gergo poco appropriato<sup>39</sup>.

Nel *rule-making*, le nuove tecnologie consentono inoltre la traduzione in codici di una regola esistente, la scrittura in tale forma di nuove regole o la riforma di queste. Le norme potrebbero poi essere scritte *ab origine* in doppio formato, tradizionale e in codice, in attuazione del più generale principio del *digital by default*<sup>40</sup>. Ciò consente di facilitare l'attività di *drafting*, supportando, ad esempio, la rilevazione delle incongruenze o incompatibilità tra norme<sup>41</sup>. Al contempo, il "rules as code" agevola l'implementazione, consentendo l'automazione dell'attuazione amministrativa che non implichi discrezionalità<sup>42</sup>, come sembrerebbe possibile anche nel nostro ordinamento giuridico<sup>43</sup>.

<sup>36</sup> Administrative Conference of the United States, Recommendation 2021-1, *Managing Mass, Computer-Generated, and Falsely Attributed Comments*, 17 giugno 2021.

<sup>37</sup> Al riguardo la Commissione europea e la Corte dei Conti suggeriscono di effettuare un'analisi separata dei commenti di massa e di evidenziarne a parte i risultati (Better Regulation Toolbox 2021, 75, 472-474, Court of Auditors, *'Have your say!': Commission public consultations engage citizens, but fall short of outreach activities*, special report 14/2019, 39-40). Sul punto si vedano S.J. BALLA, A.R. BECK, E. MEEHAN, A. PRASAD, *Lost in the flood?: Agency responsiveness to mass comment campaigns in administrative rulemaking*, in *Regulation & Governance*, 2020 e la series of essays, *Mass Comments in Administrative Rulemaking*, in *The Regulatory Review*, 13-21 dicembre 2021 (<https://www.theregreview.org/2021/12/13/mass-comments/> ultima consultazione 19/06/2022); S. KATZEN, *Public input in rulemaking*, in *The Regulatory Review*, 7 marzo 2022 (<https://www.theregreview.org/2022/03/07/katzen-public-input-in-rulemaking/> ultima consultazione 19/06/2022).

<sup>38</sup> L'impiego dell'IA nell'attività di CONSOB, AGCOM e ARERA, cit., 215 ss.

<sup>39</sup> «Addressing the lack of participation by marginalized communities in regulatory decision-making is crucial, but there is another fundamental issue. The input of marginalized communities will not matter if agencies ignore or devalue it because these insights are not expressed using the standard narratives of policymaking» (S.A. SHAPIRO, *Marginalized Groups and the Multiple Languages of Regulatory Decision-Making*, in *The Regulatory Review*, 14 marzo 2022, <https://www.theregreview.org/2022/03/14/shapiro-marginalized-groups-multiple-languages/> ultima consultazione 19/06/2022).

<sup>40</sup> Comunicazione della Commissione europea, *Bussola per il digitale 2030: il modello europeo per il decennio digitale*, COM (2021)118 final e Better Regulation Toolbox 2021, TOOL#18, 145 e TOOL#28, 228 ss.

<sup>41</sup> «The idea with "rules as code" is that the government would make its single coded version available via an API [Application Programming Interface] to the public, including developers, not just those in government» (M. WADDINGTON, *Machine-consumable legislation: A legislative drafter's perspective – human v artificial intelligence*, in *The Loophole*, giugno 2019, 27).

<sup>42</sup> «Objective criteria in the legislation are a prerequisite for automation of case processing. [...] For example, the criterion "majority of the year" may be subject to discretion and interpretation whereas "more than 250 calendar days per year" can be assessed objectively. If fully automated case processing is introduced based on objective criteria» (così il documento dell'agenzia danese per la digitalizzazione del 2018 *Guidance on digital-ready legislation - on incorporating digitisation and implementation in the preparation of legislation*, 11, [https://en.digst.dk/media/20206/en\\_guidance-regarding-digital-ready-legislation-2018.pdf](https://en.digst.dk/media/20206/en_guidance-regarding-digital-ready-legislation-2018.pdf) ultima consultazione 19/06/2022).

<sup>43</sup> Consiglio di Stato, sez. cons. att. norm., 17 novembre 2020 n. 1322.



Non mancano però potenziali risvolti negativi. In primo luogo, la traduzione o la scrittura anche in codici comporta una estrema semplificazione (aspetto positivo)<sup>44</sup>, ma che può tradursi in un impoverimento del dettato normativo, se non in una distorsione in ipotesi estreme<sup>45</sup>. In secondo luogo, il “rule as a code”, quando porta all’automazione della decisione amministrativa, implica una tendenziale sovrapposizione tra *rule-making* e *adjudication* con perdita delle garanzie procedurali ad entrambi i livelli<sup>46</sup>. È dunque importante assicurare la trasparenza, rendendo esplicita la formulazione in codici delle regole e consentendo la partecipazione degli interessati a questo livello, così come dovrebbe essere reso pubblico il fatto che alcune previsioni non discrezionali ricevano applicazione automatizzata<sup>47</sup>.

L’intelligenza artificiale può essere utilizzata anche per adeguare le regole al mutamento del contesto o del quadro normativo di riferimento. La revisione dello *stock* può non rivelarsi una operazione neutra, se espressione di politiche de-regolatorie o al contrario favorevoli a una regolazione pervasiva, esiti particolarmente preoccupanti se influenzati da portatori di interessi di parte ben organizzati<sup>48</sup>.

In generale, trasparenza e partecipazione dovrebbero dunque essere sempre assicurati, unitamente ad un costante monitoraggio e una valutazione *ex post* dei risultati conseguiti, sia per la definizione di nuove regole, che per la revisione di quelle esistenti con l’ausilio dell’intelligenza artificiale.

#### 4. Intelligenza artificiale per l’attuazione amministrativa

Le numerose applicazioni di algoritmi lineari ai procedimenti per l’adozione di decisioni individuali (che vanno dalla mobilità dei docenti, al riconoscimento di incentivi ad imprese, ai contratti pubblici, alla determinazione di tariffe di servizi pubblici) hanno portato il giudice amministrativo italiano<sup>49</sup> a

<sup>44</sup> Ad esempio, il citato document danese suggerisce che «the rules should be worded clearly and simply, unambiguously and consistently. Simple rules do not necessarily mean a brief law text. It may require more words to make it unambiguous and clear what the rules are. This does not, however, change the overall legal principle that superfluous words in the law text should be avoided» (*Guidance on digital-ready legislation*, cit., 8 e 9). Dello stesso tenore il Better Regulation Toolbox 2021 della Commissione europea, TOOL#28 *Digital-ready policymaking*, in particolare p. 238.

<sup>45</sup> «Policy is often distorted when programmers translate it into code. [...] This is, in part, because the artificial languages intelligible to computers have a more limited vocabulary than human languages. Computer languages may be unable to capture the nuances of a particular policy. Code writers also interpret policy when they translate it from human language to computer code» (D.K. CITRON, *Technological Due Process*, *Washington Law Review*, 6, 85, 2008, 1261).

<sup>46</sup> D.K. CITRON, *Technological Due Process*, cit., 1249.

<sup>47</sup> «The decision is sufficiently transparent to enable the citizen to assess his/her avenues of complaint and it must be possible to verify the decision» (*Guidance on digital-ready legislation*, cit., 11).

<sup>48</sup> C.M. SHARKEY, *AI for retrospective review*, in *Belmont Law Review*, 8, 3, 2021, 374 ss.

<sup>49</sup> Consiglio di Stato, sez. VI, sentenze 13 dicembre 2019 n. 8472, 8473 e 8474, nonché Consiglio di Stato, sez. VI, sentenza 4 febbraio 2020, n. 881, cit. Tali pronunce fanno riferimento all’art. 22 GDPR, che sancisce il diritto per gli interessati a non essere sottoposti a una decisione basata unicamente sul trattamento automatizzato.



delineare un rafforzamento delle tutele tradizionali: trasparenza come reale conoscibilità<sup>50</sup> e spiegabilità, non esclusività della decisione algoritmica<sup>51</sup>. Si tratta di garanzie effettive?

#### 4.1. Trasparenza effettiva e spiegabilità: obiettivi raggiungibili?

Le applicazioni di intelligenza artificiale utilizzate a livello di amministrazioni centrali e locali non sono facilmente individuabili, nascoste in notizie di stampa, generiche comunicazioni istituzionali<sup>52</sup>, circolari<sup>53</sup> o altri documenti<sup>54</sup> (con l'unica eccezione, al momento in cui si scrive, del regolamento di Banca d'Italia citato in nota 22). Che si tratti del controllo del flusso di traffico, dell'accesso ai parcheggi, di polizia predittiva, di verifiche del corretto pagamento di tributi o contributi, dell'individuazione di abusi di mercato, manca una chiara indicazione dei procedimenti interessati, delle tecnologie utilizzate, dei criteri in base ai quali ci si è rivolti al mercato invece di produrre internamente l'algoritmo, delle ragioni dell'utilizzo, dell'impostazione o meno di un monitoraggio del funzionamento e dei relativi esiti<sup>55</sup>.

Questa limitata trasparenza determina ovviamente una limitata effettività delle garanzie di partecipazione: perché l'interessato dovrebbe chiedere l'accesso al codice sorgente se non sa che viene utilizzato un algoritmo? Il problema del limite alle garanzie di partecipazione che consegue alla mancanza

<sup>50</sup> «Tale conoscibilità dell'algoritmo deve essere garantita in tutti gli aspetti: dai suoi autori al procedimento usato per la sua elaborazione, al meccanismo di decisione, comprensivo delle priorità assegnate nella procedura valutativa e decisionale e dei dati selezionati come rilevanti. Ciò al fine di poter verificare che i criteri, i presupposti e gli esiti del procedimento robotizzato siano conformi alle prescrizioni e alle finalità stabilite dalla legge o dalla stessa amministrazione a monte di tale procedimento e affinché siano chiare – e conseguentemente sindacabili – le modalità e le regole in base alle quali esso è stato impostato» (Consiglio di Stato n. 8472/2019, cit.).

<sup>51</sup> Salvo quanto alle decisioni seriali e standardizzate (A. MASUCCI, *L'algoritmizzazione delle decisioni amministrative tra Regolamento europeo e leggi degli Stati membri*, in *Dir. pubblico*, 2, 2020, 945-946). Sui principi elaborati dal giudice amministrativo, si rinvia a E. CARLONI, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, in *Diritto amministrativo*, 2, 2020, 273 ss.

<sup>52</sup> Ad esempio, la direzione studi e ricerche INPS, in base al sito istituzionale, «fornisce supporto tecnico-scientifico all'elaborazione delle decisioni che l'Istituto assume nell'ambito delle proprie attività istituzionali attraverso [...] l'elaborazione di statistiche, di modelli di *data mining* e *machine learning*, anche in riferimento ai *big data*» (<https://www.inps.it/nuovoportaleinps/default.aspx?itemdir=53263> ultima consultazione 19/06/2022).

<sup>53</sup> Ad esempio, INPS, circolare n. 23/2010, *Funzione di accertamento e verifica amministrativa - Attuazione del nuovo modello organizzativo delle strutture territoriali di produzione previsto dalla circolare n. 102 del 12/08/2009* e circolare n. 23/2010, *Funzione di accertamento e verifica amministrativa - Attuazione del nuovo modello organizzativo delle strutture territoriali di produzione previsto dalla circolare n. 102 del 12/08/2009*.

<sup>54</sup> Come nel caso del cd. risparmiometro, menzionato nel Piano della performance dell'Agenzia delle entrate 2018-20, ma poi descritto solo in una decisione del Garante della Privacy del 20 luglio 2017, n. 321, *Sperimentazione di una procedura basata sull'utilizzo di informazioni fornite dall'Archivio dei rapporti finanziari e degli elementi presenti nell'Anagrafe tributaria per l'individuazione di profili di evasione rilevanti*.

<sup>55</sup> Il problema della limitata trasparenza resta attuale anche per numerose agenzie Nordamericane (D. FREEMAN ENGSTROM, D.F. HO, *Algorithmic Accountability in the Administrative State*, in *Yale Journal of Regulation*, 37, 2020, 800 ss.) e non sembra sia stato dato seguito all'«Agency Inventory of AI use Cases» richiesto dall'Executive Order n. 13960/2000 (J.F. WEAVER, *Everything Is Not Terminator. The Federal Government and Trustworthy AI*, in *Robotics, Artificial Intelligence & Law*, 4, 3, 2021, 291 ss.). «In the absence of timely action of the executive or legislative branches to establish procedures to mitigate for administrative agency AI accountability gaps (like bias) and transparency gaps (like being understandable and traceable), the judiciary may dictate such procedures via remands under the administrative record provision of the Administrative Procedure Act of 194623 as it first did fifty years ago with informal adjudications» (A.A. GAVOOR, *The impeding judicial regulation of artificial intelligence in administrative State*, in *Notre Dame Law Review*, 97, 2, 2022, 183).



di trasparenza si pone ancor prima di quello legato alla reale possibilità, per soggetti privi di particolari competenze tecniche, di comprendere il funzionamento di un codice sorgente. Questa stessa mancanza di trasparenza potrebbe impattare negativamente anche sull'esercizio delle garanzie processuali, dal momento che soggetti potenzialmente interessati ma non al corrente del ricorso all'intelligenza artificiale in un processo decisionale, non possono contestarne né l'uso, né gli esiti. La trasparenza diventa dunque cruciale per l'effettività delle garanzie procedurali e processuali. Si tratta di una trasparenza che, per assicurarne l'effettività, andrebbe realizzata su due livelli: l'informazione sulle tecnologie utilizzate e le relative funzionalità andrebbe messa a disposizione del pubblico in forma semplificata e facilmente comprensibile, con possibilità di approfondimento per gli interessati<sup>56</sup>.

Trasparenza significa anche spiegabilità, declinazione particolarmente delicata sia quando l'algoritmo viene sviluppato da unità specializzate all'interno di una pubblica amministrazione, che in caso di acquisto sul mercato. Se nella prima ipotesi un adeguato disegno organizzativo interno dovrebbe assicurare il costante e reciproco coinvolgimento di sviluppatori e funzionari che faranno uso del sistema (non solo per assicurare una progettazione funzionale alle esigenze dei secondi, ma anche per consentire a questi di comprendere e saper spiegare il funzionamento), nel secondo è cruciale assicurare trasparenza e verificabilità dell'operatività del sistema di intelligenza artificiale attraverso il bando di gara<sup>57</sup>. Questa è la sede sia per prevedere che il funzionamento dell'algoritmo sia oggetto di audit e "spiegabile" al soggetto pubblico (che a sua volta dovrà poterlo spiegare agli interessati), che per imporre un sistema di gestione del rischio<sup>58</sup>.

Il problema della spiegabilità si pone però anche su un piano più generale e deriva dalle caratteristiche stesse delle tecnologie *data driven*, che si basano sull'osservazione di correlazione statistiche tra dati e sull'apprendimento automatico da questi (e non sul tradizionale metodo deduttivo). A ciò si aggiunga il frequente e già menzionato ricorso a "data from the wild" (si pensi all'uso di dati raccolti dai *social*

---

<sup>56</sup> Ad esempio, il registro degli usi di intelligenza artificiale messo a disposizione del pubblico della città di Helsinki offre informazioni approfondite, pur essendo al contempo perfettamente comprensibili anche da non esperti, su: *dataset* (vale a dire le fonti di dati utilizzate nello sviluppo e nell'uso del sistema, il loro contenuto e i metodi di utilizzo), *data processing* (la logica operativa dell'elaborazione automatica dei dati e del ragionamento eseguito dal sistema e i modelli utilizzati), non discriminazione (connesso, per esempio, al numero di lingue disponibili), supervisione umana durante l'uso del servizio, *risk management* (vale a dire ai rischi legati al sistema e al suo uso e i loro metodi di gestione).

<sup>57</sup> «Establishing contractual pre-conditions for acquiring algorithmic systems ensures that systems that do not comply with specific conditions of transparency or fairness are not acquired or used by governments, or that, if a vendor fails to meet contractual conditions, they are subject to contractual liability. Procurement conditions also allow for interventions in the design of algorithmic systems, as well as during their use» (Ada Lovelace Institute, AI Now Institute and Open Government Partnership, *Algorithmic Accountability for the Public Sector*, 2021, 33-35 e 44-45, <https://www.opengovpartnership.org/documents/%20algorithmic-accountability-public-sector/>, ultima consultazione 19/06/2022).

<sup>58</sup> Linee guida per la redazione dei bandi sono state definite dal governo inglese (Guidelines for AI procurement 2020, in <https://www.gov.uk/government/publications/guidelines-for-ai-procurement>, ultima consultazione 19/06/2022) e dalla città di Amsterdam (*Standard Clauses for Procurement of Trustworthy Algorithmic Systems*, 2020 (<https://www.amsterdam.nl/innovatie/digitalisering-technologie/contractual-terms-for-algorithms/>, ultima consultazione 19/06/2022)). Il Canada ha invece optato per una lista di fornitori pre-selezionati anche in base al rispetto di "demonstrated competence in AI ethics" (<https://buyandsell.gc.ca/procurement-data/tender-notice/PW-EE-017-34526>, ultima consultazione 19/06/2022).

*media*)<sup>59</sup>, tanto che diventa arduo rendere spiegabile l'intelligenza artificiale che faccia uso del *machine learning* e praticamente impossibile quanto al *deep learning*. Un tentativo di spiegabilità dei risultati prodotti dal *machine learning* va comunque impostata già in sede di addestramento di un modello al fine, quantomeno, di rendere evidente graficamente l'importanza delle diverse variabili di input<sup>60</sup>.

#### 4.2. Non esclusività della decisione algoritmica: criticità

La non esclusività della decisione algoritmica (o sorveglianza umana nel linguaggio europeo) significa che occorre garantire l'intervento umano quando la decisione sia frutto di discrezionalità (in base alle pronunce del Consiglio di Stato), così come per i sistemi di intelligenza artificiale classificati come ad alto rischio (secondo la proposta di regolamento europeo)<sup>61</sup>.

Si tratta di un presidio fondamentale, che resta però indeterminato nella sua dimensione attuativa<sup>62</sup>. Gli interrogativi che si pongono sono numerosi, tra questi, non solo quanto intervento umano vada assicurato<sup>63</sup>, ma soprattutto come evitare che il decisore finale si appiattisca sulle risultanze istruttorie elaborate da sistemi di intelligenza artificiale, che si tratti di un funzionario richiesto di una licenza o di un ispettore chiamato a valutare quale impresa sottoporre a controllo.

<sup>59</sup> Ad esempio, N. CRISTIANINI (*op. cit.*, 23) evidenzia che «when we communicate with other users on social networks, or we access databases of content, it would be useful to drop the pretense that this is a direct interaction, making instead explicit the presence of an intelligent agent acting as intermediary- so that we can explicitly decide which engagement actions are communicative acts aimed at the other users, and which ones are aimed at the recommending agent».

<sup>60</sup> «One family of approaches fundamentally attempts to describe how important different input variables are to the resulting predictions. Within this family, some methods operate on a global, or “algorithmwide,” level; they do not ask how important certain input variables are to generating a prediction for a given individual on which a running model is deployed, but how important they were to the algorithm’s accuracy during training across many individuals. The output of such methods is known as a variable importance plot, which displays graphically the relative importances of the different input variables. [...] The other kind of importance-measuring methods attempts to operate on the individual level, explaining what the most important variables were for a given individual’s predictions. But these methods are particularly novel and have yet to be thoroughly tested» (D. LEHR, P. OHM, *op. cit.*, 708-709).

<sup>61</sup> Ad esempio, «i sistemi di IA destinati a essere utilizzati dalle autorità pubbliche o per conto di autorità pubbliche per valutare l'ammissibilità delle persone fisiche alle prestazioni e ai servizi di assistenza pubblica, nonché per concedere, ridurre, revocare o recuperare tali prestazioni e servizi», «i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per effettuare valutazioni individuali dei rischi delle persone fisiche al fine di determinare il rischio di reato o recidiva in relazione a una persona fisica o il rischio per vittime potenziali di reati», «i sistemi di IA destinati a essere utilizzati dalle autorità pubbliche competenti per verificare l'autenticità dei documenti di viaggio», «i sistemi di IA destinati ad assistere le autorità pubbliche competenti nell'esame delle domande di asilo, di visto e di permesso di soggiorno e dei relativi reclami per quanto riguarda l'ammissibilità delle persone fisiche che richiedono tale status» (allegato II alla proposta di regolamento europeo).

<sup>62</sup> B. MARCHETTI, *La garanzia dello human in the loop alla prova della decisione amministrativa algoritmica*, in *BioLaw Journal*, 2, 2021, 367 ss.

<sup>63</sup> Il giudice amministrativo italiano non lo spiega e la definizione che ne dà il Gruppo di esperti indipendenti ad alto livello sull'intelligenza artificiale istituito dalla Commissione europea è alquanto vaga: “la sorveglianza [umana] può avvenire mediante meccanismi di governance che consentano un approccio con intervento umano (human-in-the-loop-HITL), [vale a dire] «la possibilità di intervento umano in ogni ciclo decisionale del sistema, che in molti casi non è né possibile né auspicabile» (*Orientamenti etici per una IA affidabile*, Bruxelles, 2019, 18).



L'intervento umano è utile se effettivamente in grado di valutare l'adeguatezza delle indicazioni, se non di realizzare una "sorveglianza umana" sul processo che ha portato a queste. Stante la difficoltà oggettiva di quanto sembra richiedersi ai funzionari, la proposta di regolamento europeo fa ricadere sugli sviluppatori l'obbligo di rendere possibile la "sorveglianza umana" per i sistemi di intelligenza artificiale ad alto rischio<sup>64</sup>. Ma anche in questi termini appare un obiettivo di difficile realizzazione, almeno per gli stessi motivi per cui è difficile assicurare la spiegabilità quando siano in uso tecnologie di *deep learning*.

Il presidio dell'intervento umano rischia peraltro di restare lettera morta anche per motivi che nulla hanno a che fare con la tecnologia, quanto con il funzionamento del cervello umano. Sono ipotizzabili, a fronte della limitata razionalità degli individui, due possibili opposte reazioni dell'intelligenza umana di fronte ai dati derivanti dall'intelligenza artificiale: da un lato una eccessiva deferenza<sup>65</sup>, dall'altro diffidenza se non rifiuto. A falsare questo rapporto entrano in gioco veri e propri *bias* cognitivi<sup>66</sup>, come l'*automation bias*<sup>67</sup> (menzionato anche dalla proposta di regolamento europeo)<sup>68</sup> che comporta un completo affidamento nei confronti di una presunta oggettività dell'indicazione derivante dall'intelligenza artificiale. Se si dovesse verificare l'operatività di questo *bias*, la differenza tra un sistema completamente automatizzato e uno in cui ha un ruolo l'elemento umano si perderebbe completamente<sup>69</sup>. È anche possibile che si manifesti un atteggiamento opposto di rifiuto, che può derivare da sfiducia nella tecnica (*algorithm aversion*)<sup>70</sup> o da limitate competenze tecnologiche<sup>71</sup>, oppure da un vero e proprio *bias* di avversione all'algoritmo. Potrebbe, inoltre, entrare in gioco quella che è stata definita

<sup>64</sup> L'art. 14, comma 3, della proposta di regolamento europeo (COM(2021)206 def.).

<sup>65</sup> D.S. RUBENSTEIN, *Acquiring Ethical AI*, in *Florida Law Review*, 73, 2021, 797 ss.

<sup>66</sup> N. RANGONE, *Le pubbliche amministrazioni italiane alla prova dell'intelligenza artificiale: problemi e prospettive* in *Studi parlamentari e di politica costituzionale*, 209, 2021, 22-25. ivi ampi riferimenti dottrinari ai suddetti *bias*.

<sup>67</sup> L'*automation bias* viene descritto come "the use of automation as a heuristic replacement for vigilant information seeking and processing" (L.J. SKITKA ET AL., *Automation Bias and Errors: Are Crews Better Than Individuals?*, in *The International Journal of Aviation Psychology*, 2000, 10, 1, 85). Di particolare interesse a questo riguardo sono gli studi e gli esperimenti svolti con riferimento alle reazioni alle indicazioni delle "macchine" dei piloti di aereo (K.L. MOSIER ET AL., *Automation Bias: Decision Making and Performance in High-Tech Cockpits*, in *International Journal of Aviation Psychology*, 8, 1, 1997, 47 ss.) e in medicina (K. GODDARD, A. ROUDSARI, J.C. WYATT, *Automation bias: a systematic review of frequency, effect mediators, and mitigators*, in *Journal of the American Medical Informatics Association*, 19, 2021, 121 ss.).

<sup>68</sup> Così l'art. 14 comma 4, lett. a) e d) della proposta di regolamento europeo, già citata, che evidenzia anche che vanno predisposte misure affinché le persone cui è affidata al sorveglianza umana siano consapevoli «della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output prodotto da un sistema di IA ad alto rischio ("distorsione dell'automazione"), in particolare per i sistemi di IA ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche» (lett. b).

<sup>69</sup> «We are worried that if we simply thrust the human at the output end of the running model, there is very little she can do to root out bias. The human becomes a rubber stamp for the machine, providing nothing more than a cosmetic reason to lull ourselves into feeling better about the results» (D. LEHR, P. OHM, *op. cit.*, 716). Sul punto anche D.K. CITRON (*op. cit.*, 1272) che evidenzia anche come «over time, human operators may lose the skills that would allow them to check a computer's recommendation».

<sup>70</sup> B.J. DIETVORST, J.P. SIMMONS, C. MASSEY, *Algorithm aversion: People erroneously avoid algorithms after seeing them err*, in *Journal of Experimental Psychology: General*, 144, 1, 2015, 114 ss., 10-11 in Scholarly Commons.

<sup>71</sup> F. DE LEONARDIS, *Big Data, decisioni amministrative e "povertà" di risorse della pubblica amministrazione*, in E. CALZOLAIO (a cura di), *La decisione nel prisma dell'intelligenza artificiale*, Milano, 2020, 159.



*illusion of validity*<sup>72</sup> (un'ingiustificata fiducia nel ragionamento umano che spesso caratterizza gli esperti)<sup>73</sup> o il *confirmation bias*<sup>74</sup>, che induce a tener conto dei dati elaborati dall'intelligenza artificiale quando a supporto dell'intuizione o convinzione formatasi dal funzionario ed a rifiutarli in caso di difformità.

Seppure non sia possibile dare indicazioni definitive e generali rispetto al ruolo dei *bias* nel processo decisionale del funzionario, non vi è dubbio che queste diverse eventualità vadano analizzate e affrontate con strumenti adeguati. Una volta constato il potenziale ruolo di determinati *bias* (secondo alcuni nell'ambito della valutazione di impatto dell'algoritmo)<sup>75</sup> le diverse reazioni potrebbero essere affrontate con strumenti di *empowerment* cognitivo<sup>76</sup>, per aumentare la consapevolezza della loro esistenza. Ciò potrebbe avvenire con una formazione mirata sui *bias* che intercorrono più frequentemente nel rapporto uomo-macchina<sup>77</sup>. Inoltre, un sistema di *accountability* interna o esterna che porti il funzionario a individuare e indicare i dati generati dalla tecnologia che lo hanno indotto ad adottare una determinata decisione potrebbe aumentarne la consapevolezza<sup>78</sup>. Infine, un ruolo importante hanno il monitoraggio e la valutazione *ex post* degli esiti dell'uso dell'intelligenza artificiale alla luce di diversi parametri, tra i quali va inclusa l'accettazione nel sistema organizzativo.

## 5. Esigenza di un quadro normativo minimo, da specificare a livello di singola amministrazione

Come già evidenziato, le amministrazioni italiane sperimentano e utilizzano l'intelligenza artificiale per esigenze conoscitive, a supporto di servizi al pubblico, nei processi decisionali per l'adozione di *policies*, regole, decisioni amministrative, così come nei controlli. Numerosi sono i vantaggi

<sup>72</sup> A. TVERSKY, D. KAHNEMAN, *Judgment under Uncertainty: Heuristics and Biases*, in *Science*, 185, 4157, 1974, 1126.

<sup>73</sup> D. KAHNEMAN, G. KLEIN, *Conditions for Intuitive Expertise. A Failure to Disagree*, in *American Psychologist*, 2009, 517.

<sup>74</sup> Si tratta di un *bias* molto diffuso, che interessa gli individui in quanto destinatari delle regole e in quanto regolatori. Sul punto la letteratura è diffusa, si veda, per tutti, C. TAYLOR, *Biased Assimilation: Effects of Assumptions and Expectations on the Interpretation of New Evidence*, in *Social and Personality Psychology Compass*, 5, 3, 2009, 827 ss.; E. ZAMIR, D. TEICHMAN, *Behavioural law and economics*, Oxford, 2018, 399; S. STERN, *Cognitive Consistency: Theory Maintenance and Administrative Rulemaking*, in *University of Pittsburgh Law Review*, 63, 2002, 589-591.

<sup>75</sup> Per un'analisi di impatto che cerchi di individuare specificatamente queste reazioni si veda l'art. 6 della proposta di *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration* elaborate dall'European Law Institute nel 2022: «an assessment of the specific and systemic impact of the system on: [...] iii. the administrative authority itself, in particular the estimated acceptance of the system and its decisions by the staff, the risks of over- or under-reliance on the system by the staff, the level of digital literacy, and specific technical skills within the authority» (consultabile in: [https://www.europeanlawinstitute.eu/news-events/upcoming-events/events-sync/news/eli-issues-guidance-on-the-use-of-algorithmic-decision-making-systems-by-public-administration/?tx\\_news\\_pi1%5Bcontroller%5D=News&tx\\_news\\_pi1%5Baction%5D=detail&cHash=f4a2a4a677e3dcf6e391d9f0a2a9bd6a](https://www.europeanlawinstitute.eu/news-events/upcoming-events/events-sync/news/eli-issues-guidance-on-the-use-of-algorithmic-decision-making-systems-by-public-administration/?tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Baction%5D=detail&cHash=f4a2a4a677e3dcf6e391d9f0a2a9bd6a)).

<sup>76</sup> Sulla distinzione tra *nudging* and *empowerment* sia consentito rinviare a F. DI PORTO, N. RANGONE, *Behavioural Sciences in Practice: Lessons for EU Policymakers*, in A. ALEMANNI, A.-L. SIBONY (a cura di), *Nudge and the Law: A European Perspective?*, Oxford, 2015, 29.

<sup>77</sup> D.K. CITRON, *op. cit.*, 1306-1307.

<sup>78</sup> M. HALLSWORTH, M. EGAN, J. RUTTER, J. MCCRAE, *Behavioural Government. Using behavioural science to improve how governments make decisions*, The Behavioural Insights Team, 2018.

potenzialmente ricavabili, sia in termini di risparmi di risorse economiche ed umane, che di interventi più mirati anche in termini prospettici. Dunque si può dire che le nuove tecnologie supportino il diritto a una buona amministrazione.

Non mancano però i rischi, in parte gestibili, a condizione che se ne riconosca l'esistenza e se affrontati nelle sedi adeguate, senza adagiarsi su facili soluzioni formali (come il riconoscimento del diritto di accesso esteso al codice sorgente dell'algoritmo, in quanto soluzione generalizzata anche per le tecnologie più avanzate). Vi è in particolare da chiedersi se le tutele tradizionali, seppure rafforzate, possano ritenersi sufficienti, o se non debba immaginarsi un nuovo diritto amministrativo dello Stato digitale<sup>79</sup>.

Ragionare in questa direzione significa, ad avviso di chi scrive, intervenire su più livelli, uno normativo generale e uno di singola amministrazione.

Un inquadramento normativo generale minimo dell'uso dell'intelligenza artificiale nei processi decisionali pubblici<sup>80</sup> appare necessario sia per individuare le ipotesi in cui si possa ricorrere a decisioni automatizzate (previsione, per il vero, che discenderà e sarà riconducibile al regolamento europeo una volta approvato), sia per sancire regole quadro volte ad assicurare trasparenza, motivazione (dunque anche spiegabilità) e verificabilità dei sistemi ed esplicitare le regole minime da rispettare per rendere effettivi tali principi.

La trasparenza degli utilizzi potrebbe, in particolare, essere realizzata sia attraverso i siti delle singole amministrazioni, sia con un *repository* centrale (strumenti non necessariamente alternativi) che, con funzione di *single access point*, dia informazioni sui sistemi di intelligenza artificiale in uso, sui motivi dell'utilizzo e i relativi rischi, su due livelli di approfondimento. Quanto alle modalità di declinazione della trasparenza, potrebbe essere impostato un approccio su due livelli, un primo molto semplice e intuitivo, un secondo più dettagliato e da specificare in sede di disciplina di singola amministrazione, che comprenda i dati utilizzati, la logica che ne è alla base, i rischi correlati e i sistemi di gestione del rischio.

La motivazione delle regole e delle decisioni in cui abbia avuto un ruolo l'intelligenza artificiale dovrebbe essere arricchita con l'indicazione (sintetica) delle fasi di utilizzo, le caratteristiche dei sistemi di intelligenza artificiale utilizzati, le ragioni dell'utilizzo, i dati lavorati e i meccanismi predisposti per la supervisione del relativo funzionamento.

Tale inquadramento generale dovrebbe, poi, prevedere che vengano indicati (a livello di singola amministrazione) i presidi da predisporre per la verificabilità dei sistemi (come audit, monitoraggio e valutazione *ex post*) e le clausole minime da inserire nei bandi di gara quando la tecnologia sia acquistata sul mercato.

Spetterà, inoltre, alla disciplina delle singole amministrazioni indicare, eventualmente, chi decide del ricorso a nuove tecnologie, se avviare una sperimentazione interna o acquistare sul mercato, gli indicatori per il monitoraggio e la valutazione *ex post* degli esiti dell'uso dell'intelligenza artificiale nei processi decisionali alla luce di parametri rilevanti per l'amministrazione di riferimento, ma che

---

<sup>79</sup> L. TORCHIA, *Lo Stato digitale e il diritto amministrativo*, in *Liber amicorum*, dedicato a M. D'Alberti, Torino, in corso di pubblicazione.

<sup>80</sup> S. CIVITARESE MATTEUCCI, *Public Administration Algorithm Decision-Making and the Rule of Law*, in *European Public Law*, 27, 1, 2021, 103 ss.

Law  
NIS

dovrebbero comunque comprendere la rilevazione degli errori riscontrati<sup>81</sup> e un'analisi controfattuale ove possibile<sup>82</sup>. In altre parole, la decisione di fare ricorso all'intelligenza artificiale o all'intelligenza umana supportata dall'intelligenza artificiale non andrebbe presa una volta per tutte, ma dovrebbe essere oggetto di valutazione alla luce dei risultati raggiunti<sup>83</sup>.

---

<sup>81</sup> R. CAVALLO PERIN, I. ALBERTI, *Atti e procedimenti amministrativi digitali*, in R. CAVALLO PERIN, D.U. GALETTA, *Il diritto dell'amministrazione pubblica digitale*, Giappichelli, 2020, 153. D. FREEMAN ENGSTROM, D.F. HO (*op. cit.*) propongono di impostare un "prospective benchmarking", vale a dire una comparazione tra casi decisi con il supporto dell'intelligenza artificiale e frutto di un processo decisionale tradizionale «This "human alongside the loop" approach provides critical information and a comparison set to help smoke out when an algorithm has gone astray, when encoding the past may miss new trends, when an algorithm may create disparate impact, or when "automation bias" causes excessive deference to machine outputs».

<sup>82</sup> «Regulators should develop tests for considering when the outcomes an algorithm creates are impermissible, based on regulatory policy goals. Regulators should begin by asking meaningful questions that can be answered by examining algorithmic outcomes, such as whether similarly situated borrowers are treated differently or whether the move from traditional pricing to algorithmic pricing has increased disparities» (T.B. GILLIS, *op. cit.*).

<sup>83</sup> C. COGLIANESE, A. LAI, *Digital Versus Human Algorithms*, in *Duke Law Journal*, 71, 2022, 1281 ss.