Intelligenza artificiale, poteri pubblici e rule of law

18 Ottobre 2023

di **Barbara Marchetti**

La portata *disruptive* e pervasiva dell'intelligenza artificiale (IA) nella nostra società, nei processi economici e di lavoro, nella politica e nell'operato delle amministrazioni pubbliche è a tutti nota e costituisce ormai un fenomeno universalmente analizzato e dibattuto. Dell'intelligenza artificiale vengono opportunamente richiamati gli innumerevoli vantaggi, ad esempio per la ricerca farmacologica, la diagnostica, la lotta al *climate change*, ma anche costi energetici e rischi in termini di privacy, sicurezza e diritti fondamentali. Qui si intende toccare, tuttavia, un profilo specifico, quello dell'uso dell'IA da parte delle amministrazioni e delle sue implicazioni per la *rule of law*.

In tutto il mondo, le amministrazioni stanno sperimentando ed utilizzando sistemi di IA in misura via via crescente. Nel 2015 esistevano in Europa 5 sole applicazioni di IA nella sfera pubblica; nel 2021 queste sono divenute 161 e nel 2022 se ne potevano contare 868. La loro crescita è esponenziale. Fino ad ora solo il 2% di tali sistemi è impiegato per l'adozione di decisioni individuali, ma si può presumere che anche tale uso crescerà significativamente.

L'impiego di IA da parte delle amministrazioni avviene in un ambiente privo di regole giuridiche: al momento, in Italia, mancano sia norme generali che linee guida, ad eccezione dell'art. 30 del nuovo codice appalti. La Commissione europea, come è noto, ha pubblicato una proposta di regolamento nel 2021 che attende ancora di essere approvata, ma in ogni caso la sua applicazione non potrà aversi prima del 2026. Nel frattempo, è bene chiedersi cosa possono fare gli Stati e quali regole potrebbero essere adottate, e domandarsi se la regolazione prevista dall'Unione, applicabile tanto agli attori privati quanto alle amministrazioni, assicuri il rispetto dei *public values* e fornisca adeguate garanzie ai cittadini.

Differenti sistemi di IA, differenti usi, differenti trade-off.

Per rispondere agli interrogativi appena mossi occorre considerare rischi e opportunità legati ai sistemi di IA, anche in relazione agli impieghi che di essi vengono fatti all'interno delle pubbliche amministrazioni. Prima di tutto, occorre considerare che i sistemi di IA basati su algoritmi deterministici (che lavorano con *hard rules*, i c.d. sistemi esperti) sono da tempo in uso nelle p.a. e non hanno un impatto così dirompente nelle relazioni tra amministrazione e cittadini. Pongono sì un problema di comprensibilità e accessibilità, ovvero di traduzione dalla regola tecnica in regola giuridica, come ha osservato il Consiglio di Stato (2270/2019; 881/2020), ma il loro funzionamento è improntato alla logica *if-then*: è la p.a. che decide le istruzioni da dare alla macchina, e quest'ultima giunge agli *output* attraverso passaggi predefiniti e predeterminati dal programmatore, spiegabili e ripercorribili a ritroso. I sistemi esperti non sono tuttavia capaci di elevate performance, sicché da tempo l'amministrazione sta sperimentando ed utilizzando algoritmi più complessi e sofisticati che auto-apprendono dall'esperienza (c.d. *machine learning-ML* e *deep learning-DL*) e che lavorano secondo connessioni che vorrebbero rifarsi a quelle delle reti neurali cerebrali umane. I risultati che questi sistemi possono assicurare sono sorprendenti: sono in grado di estrarre conoscenza da enormi mole di dati e tradurli in predizioni, raccomandazioni e decisioni che sono l'esito di logiche di inferenza statistico-probabilistica. Il problema però è che né il programmatore, né la macchina sono in grado di fornire spiegazioni sul perché, dati certi *input*, il sistema sia giunto a certi *output*. Questa loro caratteristica, resa con il termine di *black box*, impedisce di spiegare le ragioni di una decisione, di rintracciarne la logica e finanche di capire quali dati, tra i molti di addestramento della macchina, sono stati considerati rilevanti per il processo (nel caso di ML non supervisionato). Se vi sono errori negli *output*, inoltre, tale opacità impedisce di risalire e correggere l'errore.

È evidente quindi che essi si pongono potenzialmente in contrasto con i principi di pubblicità e trasparenza delle decisioni pubbliche, oltre che con il principio di motivazione delle decisioni individuali dell'art. 3 della l. 241/90 e dell'art. 41 della Carta dei diritti fondamentali dell'Unione europea.

Ciò non significa escludere l'uso di questo tipo di algoritmi da parte delle autorità pubbliche. In realtà la loro problematicità può essere maggiore o minore a seconda dell'impiego che di essi viene fatto. Una cosa, infatti, è usarli per gestire una *chatbot* o per estrarre conoscenza dai dati al fine di indirizzare *policies* pubbliche (ambientali, di mobilità urbana, di sicurezza pubblica) o attività di vigilanza, ciò che presenta rischi contenuti; altra cosa è un loro uso in funzione decisoria. Stabilire l'assegnazione di una sovvenzione o l'ammissione ad una Università ricorrendo ad algoritmi di apprendimento automatico è un'operazione altamente rischiosa, sia in ragione della inspiegabilità della decisione, sia in ragione dei *bias* che possono essere contenuti nei dati di addestramento o nella programmazione del sistema. Dunque, il problema non è vietare o consentire l'uso di IA in generale, ma operare i necessari distinguo e prevedere regole adeguate a seconda delle applicazioni di cui si parla e dell'uso e delle finalità che le connotano.

La regolazione europea: una sfida complessa, una regolazione sufficiente?

Come è noto regolare l'IA – anzi le IA – è operazione molto complessa, e non a caso i Paesi protagonisti dello sviluppo e della ricerca nel settore (Stati Uniti e Cina) sono lontani, per diverse ragioni, dalla adozione di una regolazione soddisfacente. Le ragioni di tale complessità sono note: la velocità degli sviluppi tecnologici rischia di rendere la regolazione – spesso affidata alle assemblee parlamentari – obsoleta; inoltre, se i caratteri dell'IA suggeriscono un approccio precauzionale (si pensi alla IA generativa e all'allarme che essa sta generando tra i suoi stessi sviluppatori), regole troppo rigide possono essere tecnologicamente poco sostenibili e finire con l'inibire la ricerca e lo sviluppo sull'IA. Su tali difficoltà pesa, infine, la competizione tra i principali *global players*, i quali temono che una regolamentazione pesante possa tradursi in un vantaggio per il competitor e in un conseguente pericolo in termini di sicurezza nazionale. Ogni critica alla proposta europea deve dunque muovere da queste premesse.

La proposta di regolamento della Commissione del 21 aprile 2021 è oggi giunta alla fase conclusiva dell'iter legislativo, dopo un intenso dibattito con Parlamento e Consiglio che ha portato a molte proposte di emendamento al testo originario. L'atto propone una regolamentazione basata sul rischio dei sistemi di IA e distingue tra sistemi vietati perché a rischio inaccettabile (art. 5), sistemi ad alto rischio (art. 6 e all. III) e sistemi a rischio minimo o basso, soggetti solo a obblighi informativi. Ovviamente non vi è spazio in questa sede per una disamina della proposta, che consta di una ottantina di articoli: si vorrebbe però richiamare l'attenzione sui sistemi ad alto rischio, perché ad essi è dedicata la gran parte della disciplina europea e ad essi si correla maggiormente il ruolo delle amministrazioni, sol che si guardi all'elenco dei settori contemplati nell'all. III: si tratta dell'immigrazione, dell'istruzione e dell'Università, della giustizia, della salute, del credito, dei procedimenti elettorali, delle infrastrutture critiche, in cui è centrale l'azione dei pubblici poteri e fondamentale deve dunque essere il rispetto dei valori pubblici, primi fra tutti eguaglianza, pubblicità e giustiziabilità.

L'immissione in commercio dei sistemi di IA che vengono impiegati in tali ambiti è soggetta ad un controllo di conformità rispetto ai requisiti che sono stabiliti dagli artt. 10 a 14 della proposta. In particolare, in base a tali norme i sistemi devono essere sufficientemente trasparenti, basarsi su dati corretti, completi e rappresentativi, devono essere accurati e sicuri, ed essere posti sotto il controllo vigile di un umano, che non deve fare affidamento eccessivo nella macchina (c.d *human in the loop*), ma deve essere in grado di comprenderne il funzionamento, verificare e correggere gli output e anche interromperne il funzionamento, qualora ciò sia necessario per evitare danni.

In termini generali, non vi è dubbio che la disciplina rappresenti un buon compromesso sul piano della sostenibilità giuridica, economica e tecnologica delle sue regole, ma la sua effettività è tutt'altro che scontata. Alcuni rischi: la verifica di conformità è operata non da una pubblica amministrazione, ma da chi fornisce o impiega il sistema; la garanzia della sorveglianza umana richiede – per le autorità pubbliche – funzionari con competenze adeguate che nelle amministrazioni scarseggiano: d'altro canto, il piano di formazione promesso nel PNRR italiano non sembra facilmente attuabile. Gli obblighi di informazione e trasparenza stabiliti nell'art. 13 della proposta sono certamente adeguati per i sistemi di IA funzionali alla adozione di politiche pubbliche *data-driven* o a indirizzare le attività di vigilanza e controllo*,* ma non paiono assicurare garanzie equivalenti a quelle offerte dalla tradizionale motivazione nel caso in cui il sistema produca decisioni aventi un impatto individuale.

In questo quadro, gli Stati potrebbero allora approvare alcune norme specifiche per le autorità pubbliche, fissando obblighi per le p.a. di trasparenza in ordine ai sistemi impiegati e alle loro caratteristiche; bandendo il ML decisorio, fintanto che permane il limite di spiegabilità della c.d. *black box*; promuovendo la formazione dei dipendenti con adeguati fondi, sul modello statunitense dell'*AI Training Act*; e regolando adeguatamente il rapporto funzionario-macchina al fine di assicurare una sorveglianza umana effettiva.