La sfida giuridica all'Intelligenza artificiale: l'approccio locale di Barcellona e Amsterdam

di Riccardo Calvara - pubblicato su "www.irpa.eu" - Osservatorio sullo Stato digitale, 12 aprile 2023

La regolamentazione dell'Intelligenza artificiale rappresenta per molti aspetti una delle sfide giuridiche più avvincenti dei nostri tempi. L'esigenza di regolare il ricorso a sistemi di intelligenza artificiale in ragione dei rischi legati all'implementazione algoritmica è, infatti, una questione sempre più all'ordine del giorno. Si pensi, in particolar modo, alle possibili inferenze negative in materia di privacy e governance dei dati personali, o in generale in materia di protezione dei diritti fondamentali delle persone sottoposte ad una decisione algoritmica che, riproducendo logiche sperequative, potrebbe emettere output contrari al principio di uguaglianza.

Nonostante l'urgenza di partorire soluzioni adeguate e proporzionali alla rivoluzione che l'intelligenza artificiale può comportare, la partita rimane pressoché aperta per via di molteplici ragioni di complessità tecnica, giuridica e di opportunità politica che la sfida regolatoria reca con sé.

La circostanza che l'oggetto da regolare sia una tecnologia in costante evoluzione e con tratti inediti rispetto alle precedenti rende la sua disciplina estremamente difficoltosa, come si può intuire dalla mancanza di una definizione univocamente riferita e onnicomprensiva. Esistono, ad esempio, diverse tipologie di sistemi di intelligenza artificiale. Alcuni funzionano secondo *hard rules* ossia *input* ben definiti e non ambigui che, se eseguiti correttamente dalla macchina, portano ad un risultato certo e predefinito. Sono i cd. algoritmi *model based*, che per tali ragioni pongono pochi problemi. Altri sistemi di intelligenza artificiale, invece, sono a tratti incontrollabili dall'uomo, il quale fornisce un *set* di informazioni all'algoritmo che, secondo processi di autoapprendimento come il *machine learning* o il *deep learning*, è in grado di elaborare risposte non del tutto spiegabili, frutto di rielaborazioni basate su complessi calcoli statistici e di probabilità. Si tratta dei cd. sistemi esperti, operanti secondo processi decisionali che restano opachi o del tutto non prevedibili per gli stessi programmatori, contraddistinti cioè dalla c.d. *black box*. Come tali, se non correttamente verificati e corretti tali sistemi possono produrre *bias*, cioè determinare effetti discriminatori per i destinatari.

La necessità di normare l'uso dei sistemi che si servono di intelligenza artificiale si piega poi a ragioni geopolitiche. Tale tecnologia negli ultimi decenni ha avuto uno sviluppo impetuoso, mostrando grandi potenzialità e suscitando aspettative molto elevate di cui i più importanti attori economici internazionali sembrano consapevoli visti gli ingentissimi investimenti fatti e programmati. Per questa via, un approccio eccessivamente limitate o proibizionistico da parte dei governi rischierebbe di frenare e inibire la spregiudicatezza economica delle *software house* in giro per il mondo. Viceversa, una completa liberalizzazione o la mancanza di norme precipue potrebbe comportare l'uso indiscriminato, e a quel punto discriminatorio, di questi strumenti a danno di cittadini più o meno consapevoli, diminuendo la fiducia degli stessi e dei mercati.

Le prime risposte degli ordinamenti si collocano nella maggior parte su una dimensione sovranazionale, la quale risente, a sua volta, non solo della complessità di allineare i diversi

interessi degli Stati, delle imprese, e dei cittadini, ma anche delle più o meno complesse macchine normative e burocratiche di implementazione delle *policies*. Nel contesto europeo, tuttavia, è già possibile individuare un certo *legal framework* di riferimento. Le istituzioni europee negli ultimi anni si sono trovate dinanzi all'esigenza di contemperare due interessi contrapposti: favorire la crescita e lo sviluppo degli investimenti e della ricerca nel campo dell'intelligenza artificiale e garantire l'adeguamento di tali sistemi ai valori europei. È stata quindi lanciata una visione "antropocentrica" dello sviluppo delle nuove tecnologie, che poggia sull'idea per cui l'IA non rappresenti un fine, ma un mezzo messo a servizio del benessere dell'uomo. Tale scelta si colloca all'interno della Strategia europea per l'Intelligenza Artificiale a è seguita l'adozione da parte della Commissione del Libro Bianco sull'intelligenza artificiale e successivamente della Proposta di Regolamento in materia di Intelligenza artificiale.

Molti degli obiettivi e delle finalità espresse dalla regolamentazione sovranazionale, tuttavia, dovranno essere calati nel contesto locale e raccolti dalle amministrazioni, soprattutto quelle che stanno lanciando piani di investimento volti alla digitalizzazione delle proprie comunità con l'obiettivo di erogare servizi pubblici sempre più intelligenti, funzionali ed efficienti. Nell'ottica della pianificazione urbana delle cd. *Smart cities*, quindi, diverrà necessaria la raccolta e l'analisi di dati che, elaborati da e per sistemi di IA, consentiranno all'ente locale di comprendere realmente le necessità delle diverse aree cittadine, approntare gli opportuni rimedi o adottare le soluzioni più idonee a garantire un migliore e più razionale presidio amministrativo del territorio, ovvero indirizzare le scelte dell'amministrazione in modo più consapevole nello sviluppo dei propri servizi pubblici, che rappresentano il punto di contatto tra l'amministrazione e il cittadino.

Per questo motivo, nonostante l'approccio strategico complessivo sovranazionale in materia di Intelligenza artificiale sia in larga parte il più opportuno per approntare una tutela adeguata e omogenea tra gli Stati, esso risulta spesso indifferente alle questioni territoriali. Quello urbano potrebbe essere, infatti, il livello ottimale per aumentare l'*accountability* delle amministrazioni e la consapevolezza dei cittadini in merito al fenomeno. Dal punto di vista amministrativo, inoltre, potrebbe essere il livello ottimale ove lanciare iniziative e partenariati con privati che siano realmente eticamente orientati e trasparenti. In altre parole, nell'attesa che una regolamentazione sovraordinata e completa sia finalmente in vigore, le amministrazioni locali potrebbero provvedere a regolare il fenomeno in base alle proprie competenze. Il rischio che si corre, altrimenti, è quello che le comunità vengano lasciate a sé stesse, catturate dai fornitori di servizi tecnologici più forti economicamente e tecnicamente.

Le pratiche di regolamentazione locale delle nuove tecnologie assumono solitamente la forma delle "strategie cittadine" o dei "manifesti". Uno degli esempi più interessanti in questo senso è quello della città di Barcellona. Nell'aprile 2021, la città catalana ha lanciato la sua "Strategia per l'uso etico di algoritmi e dati" orientata a migliorare la coscienza pubblica in rifermento all'uso degli algoritmi nei servizi pubblici e a ridistribuire tra i cittadini il valore economico e sociale prodotto dai loro dati.

La Strategia delinea individui sette principi fondamentali che i progetti locali di IA devono rispettare: 1. Azione umana e supervisione: qualsiasi iniziativa di IA che abbia un impatto sui residenti deve essere supervisionata da parte di esseri umani, con un livello di supervisione proporzionato al rischio che la tecnologia utilizzata comporta; 2. Robustezza tecnica e sicurezza: l'ente deve essere proattivo nel garantire protezione e sicurezza dei propri sistemi, effettuando *audit* di *routine* per limitare i rischi degli attacchi informatici; 3. *Privacy* e *governance* dei dati: durante l'intero ciclo di vita dei dati, la protezione e il

mantenimento della *privacy* dei residenti deve essere garantita secondo principi e processi di minimalizzazione e anonimizzazione dei dati, che devono essere di alta qualità, rilasciati in forma aperta e, ove possibile, accompagnati da informazioni sulla provenienza e sulle strategie di mitigazione degli eventuali errori. L'uso dei dati deve *compliance* con il GDPR. 4. Trasparenza e informazione: l'intero processo decisionale del sistema di intelligenza artificiale, dal momento in cui i dati vengono etichettati al momento in cui l'algoritmo prende la sua decisione, deve essere comprensibile al pubblico. L'utilizzo di sistemi di IA devono essere comunicato al destinatario con chiarezza e i suoi creatori devono essere facilmente identificabili; 5. Diversità, inclusione ed equità: per prevenire fenomeni di discriminazione l'ente deve organizzare consultazioni pubbliche con i cittadini per comprenderne l'impatto e poter così garantire l'accessibilità digitale attraverso un "approccio di progettazione universale"; 6. Impegno sociale e ambientale: le modalità di utilizzo dell'IA devono contribuire al raggiungimento dei 17 Obiettivi di Sviluppo Sostenibile delineati dalle Nazioni Unite; 7. Responsabilità, rendicontazione e controllo democratico: quando le decisioni vengono prese da sistemi di 'IA, i cittadini oggetto delle decisioni devono essere in grado di ottenere una spiegazione comprensibile, e dovrebbe essere loro garantito di contestare la decisione.

Un altro significante esempio di gestione del rischio legato all'implementazione di sistemi di IA arriva dalla Città di Amsterdam e dalle sue Condizioni di contratto standard elaborate per la prima volta nel 2020 e poi modificate nel 2021 per renderle *compliant* rispetto al cd. AI Act dell'UE. Lo scopo di tali clausole standard contratto è quello di consentire all'ente, mediante l'autodisciplina delle procedure di acquisizione di soluzioni *software* e *hardware* che utilizzano l'Intelligenza artificiale, di adottare sistemi che rispettino i principi della cd. Intelligenza artificiale antropocentrica, che siano affidabili allo stato dell'arte e che rispondano alle esigenze dell'amministrazione e dei destinatari dell'azione amministrativa. L'articolato, tra le altre cose, si dedica a: individuare l'ambito di applicabilità delle disposizioni; declinare il concetto di qualità del dato e del sistema acquistabile; definire i diritti sui dati delle parti contraenti; determinare il concetto di qualità del sistema acquistabile; definire i concetti di trasparenza e spiegabilità; delineare la strategia di mitigazione del rischio comportato da un sistema di IA.