

L'AI Act fra complessità e responsabilità: il caso della giustizia

Giovanni Maria Flick

Presidente emerito della Corte costituzionale. Mail: segreteria@gmflick.it

Con lo sviluppo e la diffusione dei sistemi di intelligenza artificiale “di dominio pubblico” è aumentata la richiesta di regole e dunque del coinvolgimento del ruolo del diritto sotto molteplici profili per conformare l’uso dell’intelligenza artificiale ai valori fondamentali della convivenza. Dignità, autonomia e libertà delle persone, integrità e sicurezza nel rapporto fra uomo e macchina diventano fondamentali.

Esse sono a rischio in un “mondo digitale” che non si limiti a interventi di miglioramento e aiuto all’attività umana; ma miri a sostituire quell’attività. La maggiore autonomia della macchina apre la via ai problemi e ai rischi che ne nascono in tema di democrazia, di autodeterminazione, di dignità e di responsabilità.

Lo sviluppo di questa nuova realtà ha portato all’apertura di un nuovo fronte non soltanto di ordine sociale ed economico; ad esempio per il problema del lavoro e della rinnovazione nei suoi contenuti, rispetto a quelli già tradizionali e consolidati; ma anche e prima ancora a un fronte di ordine etico.

Quest’ultimo aspetto concerne in particolare le decisioni che possono essere prese dalla macchina in conseguenza del suo autoapprendimento. In sostanza concerne l’equilibrio del rapporto tra macchina e persona umana a seguito della crescita non agevolmente prevedibile delle potenzialità di essa.

Alla richiesta ormai pressante di regole – nell’alternativa tra etero e self-regulation – l’Unione Europea ha cercato di rispondere con l’AI Act da

essa approvato per prima a livello mondiale. Dopo una lunga e faticosa elaborazione ha regolato l’intelligenza artificiale con lo scopo (qualcuno pensa con l’illusione) di rispettare e tutelare i diritti fondamentali e la dignità delle persone.

L’IA viene definita «sistema basato su macchine, progettato per operare con vari livelli di autonomia che può mostrare capacità di adattamento dopo l’implementazione e che, per obiettivi espliciti o impliciti, deduce dagli input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali».

La definizione forse eccede nelle affermazioni possibiliste e dubitative. Rischia di proporre una regolamentazione troppo dettagliata e troppo complessa, di carattere burocratico e di difficile applicabilità: sia rispetto all’evoluzione continua della tecnologia; sia rispetto allo stato dei mercati che vedono come antagonisti e protagonisti gli Stati Uniti e la Cina con il loro progresso tecnico; sia infine rispetto alla possibilità che i concorrenti europei finiscano per essere esclusi dalla competizione e dalla corsa al progresso tecnologico.

L’eccesso di complessità conseguente alla molteplicità dei problemi da affrontare può spingere a una specializzazione “deresponsabilizzante” ed esasperata di fronte alla realtà e complessità della vita e dei problemi da affrontare e risolvere.

La scelta dell’Unione Europea segue una traccia già ampiamente segnata – se pure con molte difficoltà – nel percorso verso l’unità europea attraverso la ricerca del difficile equilibrio tra interessi (soprattutto economici, ma anche politici) e diritti.

Sotto questo aspetto la scelta merita condivisione e apprezzamento. Risponde a un’esigenza di etica laica – non soltanto morale e religiosa –

Fiumi

ispirata al valore della persona e alla sua tradizione consolidata nella storia europea e nel nostro passato.

La ricerca di un denominatore comune per tutte le varianti di intelligenza artificiale in continuo aumento – in svariati settori – sembra riconducibile (con una sintesi di semplificazione) a una “riserva di umanità”. Su di essa insistono gran parte degli osservatori e degli studiosi della materia.

Altro e ben più complesso problema è quello di approfondire i presupposti, i limiti, i condizionamenti, i contenuti e i modi per salvaguardare e tutelare una simile riserva.

L’impianto generale dell’AI Act sembrerebbe consentire l’utilizzo dell’intelligenza artificiale per finalità organizzative in ambito giudiziario e per il supporto al magistrato per la redazione degli atti. La nuova tecnologia potrebbe inoltre essere utilizzata per implementare gli strumenti di ricerca nelle banche dati e la raccolta dei precedenti.

È forte il rischio però che – a fronte di obblighi particolarmente articolati e incisivi – prevalga la tendenza a delegare “alla macchina” la decisione senza adempiere al dovere di controllo e di sorveglianza umani.

La macchina dovrebbe essere evidentemente priva di discrezionalità.

Il compito affidato alla macchina – alla luce dell’afflusso delle richieste e dell’entità del patrimonio informativo raccolto – è logicamente orientato verso la quantità dei casi, la velocità e l’efficienza delle risposte da elaborare; più che verso la qualità di queste ultime e la specificità dei singoli casi da decidere.

Se ci si avvia sulla strada del “giudice-robot” – o quanto meno non la si esclude – diviene molto labile la possibilità di coglierne il confine con un’azione di supporto della macchina al “giudice-persona”; e quindi di distinguere fra quest’ultimo e il robot.

Si rischia di eliminare entro certi limiti le vere e proprie “decisioni” maturate attraverso la ricerca e la riflessione sui precedenti. Si rischia di azzerare il confine tra norma giuridica e regola algoritmica; tra integrazione e conservazione di quanto è noto per l’esperienza passata e per la conoscenza acquisita dal giudice e quanto invece è stato registrato dalla macchina.

In altre parole nella ritenuta “solidità” del “referto” robotico si rischia di dimenticare i suoi limiti. Essi si riassumono nella possibilità di errori nella sua “alimentazione”; nei bias derivanti dalla sua formazione e dagli eventuali pregiudizi presenti nei dati che la macchina riceve ed elabora; nella mancanza di una vera e sicura sua neutralità, solo apparente; nella opacità del suo “ragionamento” e della “scatola nera” che lo contiene e della sua motivazione; nella sua imprevedibilità rispetto alla sua programmazione.

Sembra complesso – anche se non impossibile – ovviare a questi problemi e deficit con i rimedi proposti dai principi del processo tradizionale: il diritto di difesa; il contraddittorio tra le parti; l’obbligo di motivazione; l’impugnazione.

La “digitalizzazione della giustizia” non è una novità e porta indiscutibili vantaggi. La crisi della “giustizia umana” rafforza però la tentazione di preferire a questo punto il giudizio della macchina pur di avere un prodotto giustizia neutrale, efficiente, veloce.

Il prezzo è molto elevato. Significa rinunciare alle “riserve di umanità” della giustizia e del giudizio; all’emotività e all’empatia; al dubbio ragionevole; a una conoscenza che vada al di là dell’apparenza di una “conoscenza di tipo algoritmico”. Il prezzo dell’umanità è un requisito fondamentale della giustizia. È costituito purtroppo anche dalla possibilità dell’errore, dell’inerzia, della lentezza, degli eccessi di interpretazione, dell’arbitrio nella incontrollabilità della persona-giudice. Ma è soltanto questo che consente di



distinguere tra l'aiuto – opportuno anzi necessario – della scienza alla decisione resa da una persona e il prodotto artificiale del robot.

I limiti dell'apparente "neutralità ideale" del "ragionamento giuridico robotico" derivano dall'imprevedibilità delle conclusioni cui può giungere quest'ultimo discostandosi dagli obiettivi programmati e dai risultati perseguiti dal suo "gestore".

L'orientamento dell'Unione europea per fortuna è fondato sull'entità e sulla gradualità del rischio insito nell'uso degli strumenti tecnologici più aggiornati, perfezionati e recenti. È un uso che deve garantire in primo luogo i principi di sicurezza, di trasparenza, di tracciabilità, di non discriminazione e di rispetto dell'ambiente e dei diritti fondamentali della persona.

Per l'Unione europea è inaccettabile e vietato – per il livello elevato e inammissibile del rischio – l'uso che incida su manipolazioni cognitive, su classificazione sociale della persona, su identificazione biometrica di essa (il riconoscimento facciale da remoto in tempo reale, eccezione fatta per talune deroghe).

Sono invece considerati ad alto rischio i sistemi che influiscono negativamente sulla sicurezza e sui diritti fondamentali, con una distinzione fra i risultati che rientrano nelle direttive dell'Unione europea sulla sicurezza dei prodotti e in quelli riconducibili a otto aree specifiche (fra cui l'identificazione biometrica, il lavoro, i servizi pubblici essenziali, la migrazione).

Infine residuano gli strumenti e i sistemi a rischio minimo per la consapevolezza dell'utente e per consentirgli una decisione informata sul loro uso. Resta ferma per l'intelligenza artificiale la necessità di assicurare trasparenza dei suoi sistemi; quella di evidenziare l'origine da essa di questi ultimi; la necessità di impedirne un uso in operazioni illegali; quella, infine, di rispettare il diritto d'autore sulle fonti. In quest'ultimo aspetto si

aprono nuove prospettive al rapporto tra l'uso dell'informazione come prodotto e merce di scambio, che sottolineano una volta di più il prevalere di una logica di profitto.

Non vi sarebbe spazio così per il "giudice-robot": esso non offre le stesse garanzie di quello umano; dovrebbe essere consapevole che la "delega" alle macchine non è rinuncia alla partecipazione e alla "variante umana" delle decisioni; che il "richiamo" ai precedenti e alla loro "imponenza" non può divenire condizionamento del gregge (il cosiddetto effetto montone); e che la "dittatura del calcolo" non può essere l'unica chiave per affrontare i problemi e i limiti della vita.

Il valore della certezza in termini di "calcolabilità giuridica" è certamente positivo, ma non tiene conto necessariamente del contesto e delle situazioni concrete "fuori campo". Quel modello di certezza e di efficienza rischia di aprire la via alla deresponsabilizzazione del giudice persona: un prezzo che appare troppo elevato.