INTELLIGENZA ARTIFICIALE E DISCRIMINAZIONE

scritto da goal | 24 Maggio 2024



di Chiara Polesel*

Dopo quasi due anni di lavoro, il 13 marzo scorso il Parlamento Europeo ha approvato il Regolamento che stabilisce le regole armonizzate sull'intelligenza artificiale (cd. "*AI Act*"). Pochi giorni prima, l'11 marzo, il Consiglio Europeo aveva confermato l'accordo provvisorio sulla **Direttiva relativa al miglioramento delle condizioni di lavoro mediante piattaforme digitali.**

Questi due testi normativi rappresentano un tentativo ambizioso di bilanciare i benefici dell'innovazione tecnologica con la necessità di proteggere i diritti fondamentali e prevenire le discriminazioni, ponendo l'Unione Europea all'avanguardia nella regolamentazione etica dell'Intelligenza Artificiale (IA) a livello globale.

L'**AI Act** definisce un quadro giuridico uniforme per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di IA nell'Unione.

Il Regolamento si pone il preciso obiettivo di minimizzare "*il rischio di discriminazione algoritmica*", promuovendo invece "*la diffusione di un'intelligenza artificiale antropocentrica e affidabile*", nel rispetto dei diritti e delle libertà fondamentali sanciti dai trattati e dalla Carta dei diritti fondamentali dell'Unione europea.

È infatti noto che questa nuova tecnologia porta con sé il

rischio intrinseco di bias e discriminazioni: l'IA è alimentata da algoritmi che apprendono da enormi set di dati e dunque essa può involontariamente replicare e amplificare i pregiudizi umani, insiti nei dati da cui apprende, con conseguenze potenzialmente discriminatorie. Per tale ragione, l'AI Act precisa che l'accesso a **dati di alta qualità** svolge un ruolo essenziale nel garantire che il sistema di IA funzioni in maniera sicura e che non diventi una fonte di discriminazione vietata dal diritto dell'Unione. Adottando un approccio basato sul rischio, il Regolamento individua quindi tre livelli:

Sistemi di IA a rischio inaccettabile: Questi sistemi sono considerati una minaccia ai valori fondamentali dell'Unione e sono pertanto vietati. Tra essi figurano i sistemi predittivi usati per valutare o prevedere la probabilità che un individuo commetta un reato, unicamente sulla base di una profilazione dei tratti e delle caratteristiche della personalità, nonché i sistemi che utilizzano meccanismi di riconoscimento delle emozioni nei luoghi di lavoro e nelle scuole.

Sistemi di IA ad alto rischio: Questi sistemi hanno un potenziale impatto significativo sulla sicurezza, sui diritti fondamentali delle persone e sui valori della società e pertanto richiedono una valutazione rigorosa della conformità e l'aderenza agli standard elevati definiti dal Regolamento, prima di poter essere immessi sul mercato. Dovranno essere sviluppati utilizzando set di dati di convalida (o apprendimento) sufficientemente pertinenti, rappresentativi, privi di errori e completi. È altresì prescritto un esame sulle possibili distorsioni, suscettibili di incidere sulla salute e sulla sicurezza delle persone, di avere un impatto negativo sui diritti fondamentali o di comportare discriminazioni vietate dal diritto dell'Unione.

Sistemi di IA a rischio limitato o minimo: Questi sistemi presentano rischi trascurabili o gestibili attraverso meccanismi di trasparenza o informazione. Dovranno essere

sviluppati in modo responsabile, seguendo le migliori pratiche disponibili e rispettando i principi generali di etica e affidabilità nell'IA.

Sempre in ottica antidiscriminatoria, tra le pratiche di IA espressamente vietate figurano anche:

- l'uso di sistemi di IA ai fini della valutazione o della classificazione delle persone o di gruppi di persone sulla base del loro comportamento sociale o di caratteristiche personali o della personalità note o previste (cd. social scoring), ogniqualvolta il punteggio sociale così ottenuto determini un trattamento pregiudizievole o sfavorevole ingiustificato o sproporzionato (tranne specifiche eccezioni consentite);
- l'uso di sistemi di categorizzazione biometrica che classificano individualmente le persone fisiche sulla base dei loro dati biometrici per trarre deduzioni o inferenze in merito a razza, opinioni politiche, appartenenza sindacale, convinzioni religiose o filosofiche, vita sessuale o orientamento sessuale (tranne specifiche eccezioni consentite).

Anche la proposta di direttiva sul **lavoro mediante piattaforme digitali** affronta il tema della discriminazione e del *bias* nei sistemi di IA applicati al cd. *gig working* e mira a garantire e promuove la trasparenza, l'equità, la sorveglianza umana, la sicurezza e la responsabilità nella **gestione algoritmica del lavoro** mediante piattaforme digitali.

La proposta di direttiva stabilisce i diritti minimi che sono garantiti a tutte le persone che svolgono un lavoro mediante piattaforme digitali nell'Unione. Quanto alle misure antidiscriminatorie, la proposta muove dai seguenti principi:

Obblighi di trasparenza: Le piattaforme devono fornire ai lavoratori informazioni trasparenti e dettagliate sui sistemi decisionali automatizzati utilizzati, ivi inclusi le categorie

di dati e azioni monitorate da tali sistemi, l’obiettivo del monitoraggio, i destinatari o le categorie di destinatari del monitoraggio, etc.

Protezione dei dati: Il trattamento dei dati personali da parte di una piattaforma di lavoro digitale mediante sistemi decisionali o di monitoraggio automatizzati è un tipo di trattamento che può presentare un rischio elevato per i diritti e le libertà delle persone fisiche. Per questo, la valutazione di impatto sulla protezione dei dati deve essere condotta acquisendo anche il parere dei lavoratori e dei loro rappresentanti.

Sorveglianza umana: Le piattaforme dovranno sorvegliare e, con la partecipazione dei rappresentanti dei lavoratori, effettuare regolarmente una valutazione dell’impatto delle decisioni prese o sostenute dai sistemi automatizzati sui lavoratori, sulle loro condizioni di lavoro e sulla parità di trattamento sul lavoro.

Nonostante i due provvedimenti analizzati segnino una svolta importante (e coraggiosa) nel panorama normativo dell’Unione, restano però ancora molte le questioni aperte sull’utilizzo dell’IA. L’auspicio è che, con l’aiuto di soluzioni tecniche, ma anche di norme chiare e uniformi e di un controllo democratico, si possano ottenere i massimi benefici dai sistemi di IA, proteggendo, nel contempo, i diritti e le libertà fondamentali degli individui.

Avvocata in Udine