LavoroDirittiEuropa
Rivista nuova di Diritto del Lavoro

Numero 3 / 2025

Armando TURSI

Intelligenza artificiale: regole, principi, strumenti

Intelligenza artificiale: regole, principi, strumenti

Armando TURSI

Professore Ordinario di Diritto del lavoro, Università degli Studi di Milano

Questo scritto è, nel contempo, la relazione introduttiva al convegno dallo stesso titolo, svoltosi presso l'Università degli studi di Milano il 9 maggio scorso; e un ampio *position paper* introduttivo di questo *forum*.

Lo scopo perseguito è quello di mettere in evidenza un processo evolutivo della regolazione giuridica dell'IA, che si pone a ridosso di due crinali: la dialettica tra regole e principi, da un lato; la sinergia tra etica e diritto, dall'altro.

1. L'intelligenza artificiale avanza con una dinamica singolare e in un certo senso paradossale, che se per un verso segna un progresso continuo nel suo utilizzo, per l'altro determina l'apertura di nuovi campi di analisi, costringendo le diverse discipline scientifiche, e in particolare le scienze umane e quelle ingegneristico-informatiche, a proiezioni cognitive reciproche e ardite, sospinte dal progredire quasi inarrestabile delle tecnologie.

In un ordinamento come quello europeo, che ha fatto della regolazione giuridica dell'IA un'area strategica del proprio intervento, una delle aree di studio in materia di IA non poteva che essere quella giuridica. Non a caso, ad oggi, nello spazio di pochi anni, si registra un rapido avanzamento regolativo, che si snoda attraverso almeno tre filoni normativi, tutti radicati nel diritto europeo.

Il filone principale è quello che stabilisce regole armonizzate per l'uso dell'IA nell'UE, con il regolamento n. 2024/1689.

Attorno a questo nucleo centrale ruota un gruppo di norme a-sistematiche, afferenti a diversi nuclei tematici e disciplinari, quali la protezione dei dati personali (art. 22 GDP 2016/679) e la trasparenza delle condizioni di lavoro (direttiva 2024/3831/UE)[1].

Come s'è detto, il terreno sovrastante a questo corpo di norme ha già iniziato a essere dissodato, anche in questa rivista [2], e anche in chiave interdisciplinare [3]; eppure, trasportato da questa marea montante di norme, un nuovo fronte analitico si apre, caratterizzato dall'avvento di nuove tecniche normative e sanzionatorie, e perfino da un apparente arretramento delle tecniche protettive giuslavoristiche tradizionali, basate sulla norma inderogabile e, da almeno un ventennio a questa parte, sulla tutela risarcitoria dei danni non patrimoniali alla persona.

Vero è che l'IA, nel volgere di pochi anni è diventata il motore di un processo di cambiamento tecnologico, economico, sociale e giuridico, ancora agli inizi e perciò incerto nelle potenzialità, negli ambiti, negli strumenti, negli esiti; quel che è certo, però, è che si tratta di un processo pervasivo, e che ciò si riflette sulla molteplicità dei piani d'impatto, che riguarda, in primo luogo,

[1] *A. Tursi*, *Il "decreto trasparenza": profili sistematici e problematici*, *Lav. Dir. Eur.*, 3/2022
[2] *Lav. Dir. Eur.*, numero antologico speciale su Intelligenza artificiale e lavoro di oggi e di domani, 21 ottobre 2024.
[3] *L. Floridi*, *Ultraintelligent Machines, Singularity, and Other Sci-fi Distractions about AI*, *Lav. Dir. Eur.*, 3/2022.

i diversi profili delle manifestazioni fattuali ed effettuali dell'IA, e in secondo luogo la loro regolazione giuridica.

Può dirsi, anzi, che le peculiarità della regolazione giuridica dell'IA non siano minori di quelle attinenti alle sue manifestazioni fattuali, sviluppandosi esse su un terreno nuovo, inaugurato da direttive e regolamenti europei che mirano a tutelare il lavoro attraverso la proceduralizzazione e la trasparenza della gestione aziendale, quando non direttamente costruite a misura dell'intelligenza artificiale.

La complessità della regolazione giuridica dell'IA si manfesta anche sul piano diacronico, attraverso il susseguirsi in rapida sequenza, non solo di norme diverse, ma anche di tecniche e politiche legislative. In effetti, nel processo di legificazione dell'intelligenza artificiale si registra, dopo una prima fase di regolazione d'impronta prevalentemente lavoristica, dapprima una seconda fase a forte impronta di protezione della *privacy*, e infine, una terza fase, in pieno sviluppo, caratterizzata a sua volta da tre elementi: la regolazione unilaterale promanante dall'azienda; i contenuti matagiuridici tendenti a una filosoficamente problematica etica d'impresa; la stretta connessione con il management realizzata attraverso appositi comitati. Si delinea in tal modo il parziale abbandono della strumentazione lavoristica genetica di norme inderogabili legali e collettive, per approdare a regolamenti e *policies* posti a presidio dei diritti esposti a rischio dall'intelligenza artificiale.

Nel dibattito pubblico sulla regolazione ottimale dell'intelligenza artificiale sembrano contrapporsi due punti di vista: un punto di vista ottimistico, orientato a minimizzare gli ostacoli giuridici e massimizzare i vantaggi economici dell'IA; un punto di vista pessimistico, preoccupato soprattutto di tenere l'IA sotto il controllo di diritti di dignità, non discriminazione, riservatezza, privacy. In realtà, il meccanismo giuridico che permea il regolamento europeo sull'IA, si basa sull'idea che i diritti dei lavoratori siano esposti a rischi per effetto dell'utilizzo di questa tecnologia digitale altamente intrusiva nella gestione delle risorse umane, e che l'obiettivo protettivo possa essere perseguito imponendo procedure e regole finalizzate a prevenirli.

A ben vedere, i rischi giuslavoristici indotti dall'IA e "selezionati" dal regolamento UE come "elevati", sono tali perché caratterizzati da deficit comunicativi nel rapporto tra *deployer* e lavoratori, a loro volta causa ed effetto della gestione algoritmica dei rapporti di lavoro: ciò rende necessario garantirne la trasparenza attraverso diritti di informazione; di qui una certa ripetitività nella loro menzione, tra GDPP, direttiva piattaforme digitali, Regolamento UE sull'IA.

L'accento che il Reg. UE pone sui codici etici come forme di regolazione dell'IA, a ben vedere, denota una sfiducia di fondo sulla possibilità di una disciplina incrementale dei rapporti di lavoro algoritmici basata su diritti e norme inderogabili a struttura lavoristicamente corrispettiva; esso, piuttosto, rispecchia quella "trasformazione profonda e sistemica della realtà dovuta all'introduzione delle tecnologie digitali", che Luciano Floridi , includendola tra i "concetti fondativi" dell'IA, definisce come "re-ontologizzazione della realtà" .

Ad essa si accompagnano altri concetti fondativi dell'IA, tra i quali spicca la distinzione tra intelligenza e agire intelligente, la quale da ragione del fatto che l'AI è in grado di fare cose intelligenti senza essere davvero intelligente; i sistemi automatizzati agiscono come se fossero

intelligenti, ma si tratta di comportamenti operativi, non cognitivi. Attribuire loro intenzionalità o soggettività morale è un errore concettuale che può portare a distorsioni normative, anche in ambito giuslavoristico.

La sfida, dunque, non è sviluppare un'IA più potente dell'intelligenza umana (che la cd. IA è già enormemente più potente, nel suo dominio esecutivo e di calcolo, di quella umana; ma un'IA socialmente sostenibile, politicamente legittima ed ecologicamente compatibile.

Anche su questo versante, collocato al confine tra filosofia, etica e diritto, si segnala il contributo di Floridi, che individua i requisiti fondamentali di un sistema di intelligenza artificiale eticamente accettabile. Tra questi, spiccano la protezione contro la manipolazione dei predittori (non ultima, la giustizia predittiva), la spiegabilità e l'adattabilità rispetto al destinatario (che in linguaggio giuridico si traduce in un obbligo informativo se non, addirittura, in una sorta di motivazione), la non discriminatorietà.

Istruttivo e ormai diffusamente noto è il caso di una Municipalità americana che, al fine di riprogrammare la frequenza e il numero dei mezzi di trasporto pubblici, seguendo il "consiglio" dell'IA, aveva sguarnito i quartieri più popolati da neri: l'indagine condotta *a posteriori* aveva rivelato come ciò fosse dovuto al fatto che i dati utilizzati dall'IA si basavano sulle rilevazioni effettuate tramite la geolocalizzazione dei telefoni cellulari, che sono posseduti prevalentemente dai bianchi dei quartieri benestanti piuttosto che dai neri dei sobborghi urbani.

Del resto, è noto come le discriminazioni possano prescindere da un preciso intento discriminatorio, basandosi la loro illiceità sull'effetto e non sull'intenzione: è il caso delle discriminazioni "indirette", in cui un fattore apparentemente neutro produce un effetto discriminatorio che potrebbe dirsi "involontario".

La domanda fondamentale da porsi è quali strumenti e quali tecniche si prestino meglio a fronteggiare i rischi insiti nelle zone d'ombra dell'IA: quelli che non a caso vengono identificati, con un sintagma ormai diffuso, con "'opacità dell'algoritmo"; una opacità bivalente, perché riguarda sia chi utilizza l'IA per gestire un potere (tipico il caso del datore di lavoro), sia colui che ne subisce le conseguenze (tipicamente il lavoratore subordinato).

La risposta anche solo intuitiva alla domanda è che i rischi insiti nell'IA non sono fronteggiabili in maniera ottimale ricorrendo allo schema classico del diritto individuale posto da norma inderogabile, ma esigono tecniche nuove e nuovi approcci.

Quanto agli approcci, si tratta di abbracciare una visuale che vada "al di là della legge" (beyond the Law), per coprire, inevitabilmente, la sfera dell'etica.

E' solo trascendendo gli obblighi legali per attingere all'etica, che è possibile dare attuazione al principio di prudenza e prevenzione, l'unico idoneo a minimizzare i rischi (anche dell'IA). Si tratta, a ben vedere, di riattualizzare il tema della responsabilità sociale dell'impresa, assurto a una certa popolarità a cavallo tra i due millenni, e oggi riproposto in pompa magna da recenti direttive europee.[4]

[4] V. la *Corporate Sustainability Due Diligence Directive (CSDDD)* e la *Corporate Sustainability Reporting Directive (CSRD)*,

2. Tra i principi sostanziali ispiratori della regolazione etica dell'IA riveste un ruolo di assoluta preminenza il principio antropocentrico, già cardine del Regolamento europeo AI.

Il corollario di tale approccio a livello di tecnica normativa è di duplice natura: non si tratta solo di privilegiare i principi, i codici etici e le procedure rispetto alle regole, e alle norme legali; si tratta anche di ridisegnare la mappa dei poteri e delle responsabilità dell'azienda e del management, costruendo una *governance* d'impresa a misura del regolamento europeo sull'IA, come pure delle recenti direttive europee sula "dovuta diligenza" in materia di sostenibilità, e del regolamento sul bilancio di sostenibilità.

In secondo luogo, la qualità dell'impatto sull'insieme dei mondi vitali delle persone – tra i quali quello del lavoro – sembra diversa dalle "rivoluzioni" scientifiche e tecnologiche precedenti, se non altro perché si moltiplica esponenzialmente la capacità di effettuare calcoli, stime valutative, profilazioni personali, ipotesi predittive su fatti e soprattutto su condotte umane, con una precisione statistica incomparabilmente maggiore rispetto al passato, sì da dare luogo a macchine che non funzionano semplicemente in base ad algoritmi esecutivi, bensì potenziano il processo di *output* attraverso algoritmi di apprendimento o "generativi".

È questo, anzi, il dato alla base del suggestivo appellativo di "intelligenza artificiale"; ma con la fondamentale precisazione che trattasi di processi elaborativi interamente basati sulla statistica e sul calcolo di probabilità, quindi avulsi non solo da ogni conoscenza di tipo causale (in coerenza con il relativismo cognitivo dell'epistemologia moderna, da Hume in poi), ma anche da ogni intenzionalità, ponendosi per questa via in linea con le tendenze attuali dell'ordinamento giuslavoristico a coltivare la propria "vena" antidiscriminatoria (in cui, per l'appunto, almeno secondo la visione dominante, rileva l'effetto e non l'intenzionalità discriminatoria).

Il baluardo che le norme europee erigono contro i rischi di una deriva erratica (non consapevole) del processo decisionale automatizzato sta proprio nel controllo e nella supervisione di un essere umano.

Piuttosto, la moltiplicazione incrementale dei dati (*big data*) personali, programmati o auto-appresi dalla macchina, e posti alla base di decisioni interamente o parzialmente automatizzate, rende evidente come il terreno di elezione delle criticità dell'IA sia da un lato quello dei poteri datoriali (direttivo, di vigilanza e controllo), e dall'altro quello delle discriminazioni sul lavoro.

3. Anche la Direttiva sul lavoro tramite piattaforme digitali affronta il problema della opacità algoritmica prevedendo diritti di informazione e di accesso e imponendo al datore di lavoro l'obbligo di informare i lavoratori in merito ai sistemi di monitoraggio automatizzati e ai sistemi decisionali automatizzati.

Non si può non osservare, a proposito del divieto di decisioni interamente automatizzate, come esso segni un punto di netta emersione della natura non tecnico-informatica, ma cognitiva, filosofica ed etica delle questioni connesse all'IA, giacché – del tutto condivisibilmente, a nostro avviso – l'IA, concepita per sopravanzare, grazie ai *big data*, la correttezza di decisioni e valutazioni umane, viene di ritorno assoggettata allo stesso controllo umano: ciò non fa altro che confermare quanto il sintagma "intelligenza artificiale" possa essere fuorviante.

Bisogna essere consapevoli – *in primis* gli HR e i responsabili ICT delle aziende –che gli algoritmi non sono pura matematica, ma opinioni incastonate in un linguaggio matematico; la decisione di una persona non è il prodotto di un algoritmo dal funzionamento biochimico (secondo i dettami

del transumanesimo). Il problema non sta nell'umanizzazione della macchina, ma nella abdicazione decisionale dell'uomo, che può realizzarsi delegando ai *big data* e ai loro manipolatori non solo il processo decisionale ma anche la decisione.

Così facendo, per esempio, si possono produrre discriminazioni sulla base di costanti algoritmiche che generano in via automatica regole e decisioni alle quali i decisori sono abilitati a rimettersi: mentre l'atteggiamento corretto sarebbe quello di considerare l'IA alla stregua di un fonte di *non binding opinion*. Tutto ciò potrebbe apparire irrealistico e irragionevole, non essendo l'intelligenza umana in grado di processare *big data*, né di ricostruire processi basati su valutazioni probabilistiche; ma ciò che si richiede alla c.d. "persona naturale" non è questo, bensì valutare con la sua intelligenza "naturale", con il suo intuito, con la sua ragionevolezza, con il suo buon senso (tutte facoltà così poco irrazionali e concrete, oltre che esclusivamente umane, da essere poste alla base – per esempio – dei ragionamenti giuridici).

Invero, le macchine di IA non operano nel vuoto, ma si basano su algoritmi che riflettono i *biases* presenti o indotti in fase di progettazione e programmazione, o di scelta dal *dataset*, o infine di funzionamento dell'algoritmo.

Dalle considerazioni svolte emerge la ragione per cui si ritiene, almeno in Europa, che i sistemi decisionali di IA esigano una *governance* etica. In realtà, si tratta di ragioni plurime e di più profili.

In primo luogo, l'azienda potrebbe non avere le risorse per condurre un'investigazione sufficiente per adottare una decisione bene informata. Sul piano più generale, si pone la questione se sia etico produrre effetti differenziati nell'ambito di sottopopolazioni - col conseguente problema delle discriminazioni intersezionali e multiple - , o se tale rappresentazione differenziata violi il principio di eguaglianza. In terzo luogo, l'inconsapevolezza delle discriminazioni indirette rende preferibile prevenirle attraverso procedure. Infine, nel dominio della responsabilità sociale d'impresa rileva anche la dimensione supererogatoria, in cui l'imprenditore si auto-incola a condotte etiche ma non giuridicamente obbligatorie.

Né è risolutivo ricorrere alla stessa intelligenza artificiale per risolvere i problemi di equità di cui essa stessa può essere vittima o propalatrice: alla domanda "è questa una decisione equa", se sei un tecnico o uno scienziato informatico risponderai applicando una o più metriche quantitative per la misurazione dell'equità. Ciò solleva, insomma, il problema della incomprimibile pluralità delle metriche per l'equità esistenti, cha a lor volta sono incompatibili l'una con l'altra: sicché non si può essere equi alla stregua della loro totalità allo stesso tempo.

Al contempo, deve constatarsi il transito dalle regole ai principi e, sul piano sostanziale, la centralità del principio antropocentrico.

Ma, come s'è detto, l'approccio procedurale dell'AI ACT dell'UE non si esaurisce in una modifica rilevante delle norme, dalle regole ai principi; esso opera sulla natura stessa delle fonti, incentivando le organizzazioni a regolare l'Intelligenza Artificiale con un Codice Etico IA vincolante per i sistemi, che si dipani lungo i cicli gestionali e sia integrato nei processi aziendali. Sarà anche necessaria una mappatura della presenza ed interazione della Intelligenza Artificiale con il coordinamento e il presidio di un Comitato Etico.

Anche nel ddl italiano sull'intelligenza artificiale, approvato alla Camera il 25 giugno 2025 e sostanzialmente attuativo del regolamento UE, gli elementi basilari sono costituiti per un verso, sul piano sostanziale, dal principio antropocentrico e dai relativi corollari, alla cui stregua i sistemi

di intelligenza artificiale "devono essere sviluppati e applicati nel rispetto dell'autonomia e del potere decisionale dell'uomo"; e per l'altro, dal un sistema di *governance* che si segnala per la duplicazione funzionale tra due autorità nazionali: l'Agenzia per l'Italia Digitale (AgID) vocata allo sviluppo dell'IA, e l'Agenzia per la Cybersicurezza Nazionale (ACN).[5] La ripartizione di competenze così delineata risponde a una logica di specializzazione funzionale che riflette la duplice anima dell'IA, al contempo opportunità di sviluppo e potenziale fonte di rischi.

[5] *E. Guarnaccia, F. Traina, Algoritmi e diritti: il nuovo disegno di legge italiano sull'intelligenza artificiale, Lav. Dir. Eur.*, n. 2/2025.