

ISSN 1121-8762

Diritto delle Relazioni Industriali

Rivista trimestrale già diretta da
MARCO BIAGI

*Risorse umane e intelligenza artificiale
alla luce del regolamento (UE)
2024/1689, tra norme legali,
etica e codici di condotta*

Silvia Ciucciovino

anticipazione

N. 3/XXXIV - 2024

Pubblicazione Trimestrale - Poste Italiane s.p.a. - Spedizione in abbonamento postale - D.L. 353/2003 (convertito in L. 27/02/2004 n° 46) articolo 1, comma 1, DCB (VARESE)

ADAPT
www.adapt.it
UNIVERSITY PRESS

 GIUFFRÈ FRANCIS LEFEBVRE

*Intelligenza artificiale e diritto del lavoro:
problemi e prospettive*

**Risorse umane e intelligenza artificiale
alla luce del regolamento (UE) 2024/1689, tra norme legali,
etica e codici di condotta**

Silvia Ciucciovino

Sommario: **1.** La nozione giuridica di “intelligenza artificiale”. – **2.** L’IA nel campo dell’occupazione e delle risorse umane: l’impatto del regolamento (UE) 2024/1689 nella delimitazione delle ipotesi vietate, lecite e ad alto rischio. – **3.** Le applicazioni concrete della IA nel campo delle risorse umane: l’IA come strumento di lavoro e come strumento di esercizio di poteri privati. – **4.** L’approccio regolativo europeo basato sul rischio, l’*accountability* algoritmica e la responsabilità dei *deployers* (datori di lavoro, agenzie per il lavoro e altri soggetti del mercato del lavoro). – **5.** Legalità ed etica algoritmica nel rapporto di lavoro ai sensi del quadro regolativo europeo sull’IA. – **6.** Il diritto all’informazione e alla trasparenza come principio trasversale a bilanciamento del *potere tecnologico*: finzioni e funzioni della trasparenza algoritmica. – **7.** Dalla *hard law* alla *soft law* come tecnica di controllo degli impatti sociali dello sviluppo tecnologico: codici di condotta, autoregolamentazione e partecipazione sindacale. – **8.** La via italiana alla regolazione dell’IA nel campo delle risorse umane *de iure condito* e *de iure condendo*: il puzzle normativo e il rischio della *over-regulation*.

1. La nozione giuridica di “intelligenza artificiale”

Nel vasto e sconfinato panorama delle *information and communication technologies* (ICT), l’intelligenza artificiale (IA) rappresenta un sottoinsieme, ormai dotato di una precisa definizione legale a livello internazionale ed europeo. Il regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024, pubblicato sulla *GUUE* del 12 luglio 2024, definisce il sistema di IA nei seguenti termini: «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi

* *Professoressa ordinaria di Diritto del lavoro, Università degli Studi Roma Tre.*

espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali» (articolo 3, § 1, punto 1). Questa definizione ricalca quella adottata dall'OCSE nella raccomandazione sull'IA ⁽¹⁾ e ripresa dalla *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law* del 17 maggio 2024 ⁽²⁾. La convergenza tra definizione UE e definizione OCSE ha una chiara valenza politica e normativa: quella di garantire la certezza del diritto e agevolare la più ampia convergenza e accettazione internazionale delle regole in materia di sistemi di IA. D'altronde la vocazione applicativa del regolamento è a vasto raggio in quanto vincola i fornitori di sistemi di IA, a prescindere dal fatto che siano stabiliti nell'Unione o in un Paese terzo, e tutti gli utilizzatori dei sistemi di IA stabiliti nell'Unione, ma anche i fornitori e gli utilizzatori di sistemi di IA stabiliti in un Paese terzo, nella misura in cui l'output prodotto da tali sistemi sia destinato ad essere utilizzato nell'Unione (articolo 2) ⁽³⁾.

Come chiarito dal regolamento (considerando 12), esulano dalla nozione di IA accolta a livello europeo i sistemi che «utilizzano regole definite unicamente da persone fisiche per eseguire operazioni in modo automatico». L'IA quindi *non* comprende i sistemi software tradizionali o gli approcci di programmazione più semplici, ancorché integralmente automatizzati, che si basano su regole definite esclusivamente da persone fisiche per eseguire automaticamente operazioni.

Parafrasando il regolamento si può dire che una caratteristica fondamentale dei sistemi di IA è la loro *capacità inferenziale* che trascende

⁽¹⁾ *Raccomandazione sull'intelligenza artificiale: principi per la gestione responsabile di una AI affidabile e raccomandazioni agli Stati aderenti*, approvata il 3 maggio 2024 dal Consiglio dell'OCSE, che aggiorna la precedente *Raccomandazione sull'Intelligenza Artificiale* adottata nel 2019.

⁽²⁾ Sottoscritta per conto dell'UE dalla Commissione europea il 5 settembre 2024. Cfr. l'art. 2 della Convenzione, rubricato *Definition of artificial intelligence systems*, recita: «For the purposes of this Convention, “artificial intelligence system” means a machine-based system that for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment».

⁽³⁾ G. RESTA, *Cosa c'è di 'Europeo' nella Proposta di Regolamento UE sull'intelligenza artificiale?*, in *Il Diritto dell'Informazione e dell'Informatica*, 2022, n. 2, p. 323 ss.

l'elaborazione di base dei dati e consente l'apprendimento, il ragionamento o la modellizzazione. Le tecniche che consentono l'inferenza nella costruzione di un sistema di IA comprendono approcci di apprendimento automatico che imparano dai dati come conseguire determinati obiettivi e approcci basati sulla logica e sulla conoscenza che traggono deduzioni, con metodi basati sulla probabilità, dalla conoscenza codificata o dalla rappresentazione simbolica.

Va da sé, quindi, che ciò che contraddistingue strutturalmente un sistema di IA, e vale anche a distinguerlo da un generico "sistema decisionale automatizzato", è il grado di *autonomia* nella generazione di risultati (output) rispetto al coinvolgimento umano. I sistemi di IA sono progettati per funzionare con livelli di autonomia variabili e quindi sono contraddistinti da un certo grado di imprevedibilità dell'output. Si usa in proposito distinguere anche tra sistemi non deterministici e sistemi deterministici dove il fattore discriminante è dato proprio dalla imprevedibilità/prevedibilità dell'esito algoritmico ⁽⁴⁾.

Si capisce allora come i "sistemi automatizzati" non coincidono necessariamente con i sistemi di IA. Il concetto di sistema automatizzato rinvia semplicemente a sistemi di elaborazione dei dati che fanno uso di macchine (considerando 12), ma è una nozione più ampia di quella di IA e, in quanto tale, atta a comprendere sistemi digitali e software anche più tradizionali, non dotati di ragionamento autonomo rispetto all'intervento umano ⁽⁵⁾.

La capacità inferenziale, l'autonomia e la capacità di autoapprendimento (ossia la capacità di adattamento progressivo e di cambiamento autonomo durante l'uso) dei sistemi di IA rappresentano anche le caratteristiche più dirompenti dal punto di vista tecnologico e che destano preoccupazione sul piano della tutela dei diritti delle persone implicate o impattate dall'uso dei sistemi di IA ⁽⁶⁾.

⁽⁴⁾ I sistemi di IA deterministici sono caratterizzati dalla prevedibilità dell'output algoritmico. I sistemi non deterministici possono produrre risultati diversi anche con gli stessi dati di input. Questo è dovuto all'uso di algoritmi che incorporano elementi di casualità o che apprendono, evolvendosi e adattandosi nel tempo, come le reti neurali e gli algoritmi genetici.

⁽⁵⁾ Per le diverse accezioni di "automatizzazione" dei sistemi e delle c.d. decisioni algoritmiche si consenta il rinvio a S. CIUCCIOVINO, *La disciplina nazionale sulla utilizzazione della intelligenza artificiale nel rapporto di lavoro*, in *LDE*, 2024, n. 1.

⁽⁶⁾ O. POLLICINO, P. DUNN, *Intelligenza artificiale e democrazia. Opportunità e rischi di disinformazione e discriminazione*, Bocconi University Press, 2024; L. FLORIDI, *The*

Già da questi accenni si può dire che, nella prospettiva regolativa, la nozione giuridicamente rilevante di IA si rinviene ormai nelle fonti sovranazionali, che ne dettano una definizione uniforme e armonizzata a livello UE, che si sta diffondendo anche oltre i confini d'Unione. A tale nozione si deve far riferimento per identificare l'ambito applicativo della disciplina europea in materia di IA, vincolante per chi produce, chi commercializza e chi utilizza sistemi di IA.

2. L'IA nel campo dell'occupazione e delle risorse umane: l'impatto del regolamento (UE) 2024/1689 nella delimitazione delle ipotesi vietate, lecite e ad alto rischio

Guardando alle applicazioni dei sistemi di IA nel campo delle risorse umane va detto che, ai sensi del regolamento, si tratta di applicazioni di base non vietate, salvo che non siano finalizzate al riconoscimento delle emozioni sul lavoro in base a dati biometrici, cioè mirate ad inferire le emozioni di una persona fisica nell'ambito del luogo di lavoro (articolo 5, § 1, lettera *f*). Il riconoscimento delle emozioni sul luogo di lavoro, infatti, è compreso nell'elenco tassativo delle applicazioni vietate, salvo che non risponda a motivi di sicurezza. La nozione giuridicamente rilevante di "emozione" a questo scopo non appare comunque sovrapponibile a quelle caratteristiche personali e attitudinali che formano oggetto di attenzione crescente in sede di ricerca e selezione dei candidati a una posizione lavorativa e che afferiscono alle competenze c.d. non cognitive⁽⁷⁾. Queste ultime, anche dette *soft skills*, sono legate non soltanto alla dimensione relazionale e pro sociale della persona, ma anche a tratti della personalità e caratteriali (adattabilità, coscienziosità, perseveranza, senso

European Legislation on AI: a Brief Analysis of its Philosophical Approach, in *Philosophy & Technology*, 2021, vol. 34, n. 2, p. 215 ss.; L. FLORIDI, *Etica dell'intelligenza artificiale. Sviluppi, opportunità sfide*, Raffaello Cortina, 2022; P. BENANTI, *Oracoli. Tra algoretica e algocrazia*, Luca Sossella, 2018; L. FLORIDI, J. COWLS, M. BELTRAMETTI ET AL., *AI4People – An ethical framework for a good AI Society: Opportunities, risks, principles, and recommendations*, in *Minds and Machines*, 2018, vol. 28, n. 4, p. 689 ss.; M. CHIRIATTI, *Incoscienza artificiale*, LUISS University Press, 2021.

⁽⁷⁾ M. PELLEREY, *Sulle competenze, e in particolare su quelle personali dette spesso soft skill: il ruolo nel mondo del lavoro*, in *Forma@re*, 2023, n. 1, p. 5 ss.; R. CENCIOTTI, L. BORGOGNI, C. CONSIGLIO, *Le capacità agentiche e il loro ruolo nella riuscita professionale*, in *Personale e Lavoro*, 2016, n. 575, p. 13 ss.

della responsabilità, scrupolosità, amore per il lavoro ben fatto) che rientrano in un concetto ampio di attitudine⁽⁸⁾ e condizionano la capacità di agire della persona nei contesti lavorativi (*human agency* o competenze *agentiche*). Tutti fattori, questi, molto importanti, secondo la scienza pedagogica e la psicologia del lavoro, alla stregua di *mobilizzatori* delle competenze di dominio (*hard skills*) e rilevanti per il successo professionale. Non a caso, in sede selettiva, si mira a studiare questi tratti della personalità dei candidati, anche mediante l'osservazione delle loro manifestazioni esterne, ad esempio attraverso le movenze e la mimica facciale.

Per quanto ampia sia l'accezione degli elementi che concorrono a identificare le capacità non cognitive, appare sufficientemente chiaro tuttavia il loro confine rispetto alle "emozioni" la cui identificazione tramite IA nel contesto lavorativo è vietata dal regolamento. Queste ultime fanno riferimento a sensazioni quali «felicità, tristezza, rabbia, sorpresa, disgusto, imbarazzo, eccitazione, vergogna, disprezzo, soddisfazione e divertimento»⁽⁹⁾. È chiaro, però, che, se tali elementi dovessero essere presi in considerazione, anche indirettamente attraverso tecniche di approssimazione algoritmica, per identificare o inferire tratti della personalità e attitudini rilevanti ai fini lavorativi della persona, si ricadrebbe nelle pratiche vietate dal regolamento.

Visti i valori della persona in gioco, non pare, peraltro, che possa ritenersi il divieto limitato alla fase di svolgimento del rapporto escludendo

⁽⁸⁾ La componente attitudinale è definita dalla raccomandazione del Consiglio del 22 maggio 2018 relativa alle competenze chiave per l'apprendimento permanente come quella che descrive «la disposizione e la mentalità per agire o reagire a idee, persone o situazioni».

⁽⁹⁾ Aiuta a comprendere il significato dell'espressione il considerando 18 del regolamento: «La nozione di "sistema di riconoscimento delle emozioni" di cui al presente regolamento dovrebbe essere definita come un sistema di IA finalizzato a identificare o inferire emozioni o intenzioni di persone fisiche, sulla base dei loro dati biometrici. La nozione si riferisce a emozioni o intenzioni quali felicità, tristezza, rabbia, sorpresa, disgusto, imbarazzo, eccitazione, vergogna, disprezzo, soddisfazione e divertimento. Non comprende stati fisici, quali dolore o affaticamento, compresi, ad esempio, i sistemi utilizzati per rilevare lo stato di affaticamento dei piloti o dei conducenti professionisti al fine di prevenire gli incidenti. Non comprende neppure la semplice individuazione di espressioni, gesti o movimenti immediatamente evidenti, a meno che non siano utilizzati per identificare o inferire emozioni. Tali espressioni possono essere espressioni facciali di base quali un aggrottamento delle sopracciglia o un sorriso, gesti quali il movimento di mani, braccia o testa, o caratteristiche della voce di una persona, ad esempio una voce alta o un sussurro».

le applicazioni nel mercato del lavoro per le attività di *assessment e matching*, benché il regolamento riferisca il divieto all'inferenza delle emozioni «nell'ambito del luogo di lavoro» con una espressione che potrebbe effettivamente prestarsi ad una interpretazione restrittiva.

Ai sensi dell'articolo 5 del regolamento tra le applicazioni vietate rientrano anche le pratiche di punteggio sociale (*social scoring*). Si tratta di applicazioni di sistemi di IA finalizzati a valutare o classificare le persone fisiche sulla base dell'osservazione del loro comportamento sociale o di caratteristiche personali o della personalità note, inferite o previste, al fine di praticare un trattamento pregiudizievole o sfavorevole in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti oppure a praticare un trattamento sfavorevole ingiustificato o sproporzionato rispetto al comportamento sociale o alla sua gravità (articolo 5, § 1, lettera c). Non è escluso, ma si tratterebbe di casi limite, che un sistema di IA applicato alle risorse umane possa ricadere in tali ipotesi, qualora si configurino le circostanze ivi previste, ad esempio qualora si attribuisse direttamente o indirettamente una rilevanza lavorativa a pratiche di *social scoring*. C'è da chiedersi, ad esempio, se un sistema di IA che effettui una scrematura o preselezione dei candidati ad una posizione lavorativa, sulla base di informazioni relative al loro comportamento o tratti della personalità ricavate dai profili pubblici dei social network, possa configurare o meno una pratica di *social scoring* vietata ai sensi del regolamento.

Al di fuori delle ipotesi vietate citate, le applicazioni della IA nel campo delle risorse umane sono di base lecite, benché classificate, con una presunzione semplice, “ad alto rischio” ai sensi del combinato disposto dell'articolo 6, § 2, e dell'allegato III del regolamento e salve le deroghe previste dallo stesso articolo 6⁽¹⁰⁾.

L'IA può trovare numerose applicazioni, tanto nel mercato del lavoro e nel contesto delle politiche attive del lavoro, quanto nella gestione del rapporto di lavoro in tutte le sue fasi, da quella pre-assuntiva a quella gestionale fino a quella della cessazione del rapporto. Di questi molteplici potenziali usi si trova un'elencazione sintetica, ma certo non esaustiva, nell'allegato III del regolamento.

⁽¹⁰⁾ L'art. 6, § 3, prevede che, «In deroga al paragrafo 2, un sistema di IA di cui all'allegato III non è considerato ad alto rischio se non presenta un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche, anche nel senso di non influenzare materialmente il risultato del processo decisionale».

Al § 4 dell'allegato III, sotto il titolo *Occupazione, gestione dei lavoratori e accesso al lavoro autonomo*, vengono elencati i sistemi di IA destinati ad essere utilizzati: a) «per l'assunzione o la selezione di persone fisiche, in particolare per pubblicare annunci di lavoro mirati, analizzare o filtrare le candidature e valutare i candidati»; b) «per adottare decisioni riguardanti le condizioni dei rapporti di lavoro, la promozione o cessazione dei rapporti contrattuali di lavoro, per assegnare compiti sulla base del comportamento individuale o dei tratti e delle caratteristiche personali o per monitorare e valutare le prestazioni e il comportamento delle persone nell'ambito di tali rapporti di lavoro».

Al § 3, sotto il titolo *Istruzione e formazione professionale*, vengono elencati, tra gli altri e per quanto qui interessa, i sistemi di IA destinati ad essere utilizzati: a) «per determinare l'accesso, l'ammissione o l'assegnazione di persone fisiche agli istituti di istruzione e formazione professionale a tutti i livelli»; b) «per valutare i risultati dell'apprendimento, anche nei casi in cui tali risultati sono utilizzati per orientare il processo di apprendimento di persone fisiche in istituti di istruzione o formazione professionale a tutti i livelli»; c) «per valutare il livello di istruzione adeguato che una persona riceverà o a cui potrà accedere, nel contesto o all'interno di istituti di istruzione o formazione professionale a tutti i livelli».

In larga misura le applicazioni elencate dal § 4, lettera *a*, dell'allegato III sono comprese nelle attività di somministrazione, di intermediazione, di ricerca e selezione del personale, di supporto alla ricollocazione professionale, così come definite dall'articolo 2 del decreto legislativo n. 276/2003, svolte dalle agenzie del lavoro di cui all'articolo 4, nonché dagli ulteriori soggetti, compresi i siti internet, che operano in base ai regimi particolari di autorizzazione e di accreditamento di cui agli articoli 6 e 7 del decreto legislativo n. 276/2003.

Per quanto riguarda le applicazioni elencate dal § 3 è possibile ricomprendervi almeno una parte dei servizi e delle misure di politica attiva del lavoro, come definite dall'articolo 18 del decreto legislativo n. 150/2015, messe in atto dai soggetti privati accreditati dalle regioni, nonché la formazione professionale continua, ivi compresa quella finanziata dai fondi interprofessionali di cui all'articolo 118 della legge n. 388/2000, nonché dai fondi bilaterali di cui all'articolo 12, comma 4, del decreto legislativo n. 276/2003.

Chiaro è che non tutte le applicazioni di IA messe in campo dai soggetti menzionati, per quanto riconducibili alle attività di cui ai §§ 3 e 4,

rientrano necessariamente nell'elenco dei sistemi ad alto rischio di cui all'allegato III. Occorre effettuare una disamina accurata delle finalità concrete cui sono indirizzate le applicazioni di IA per comprendere se rientrano in quell'uso delineato dal regolamento come rischioso, a partire dall'esame della finalità cui il sistema è indirizzato (selezionare l'accesso alle opportunità formative, orientare i percorsi di formazione professionale e gli obiettivi di apprendimento della formazione professionale, valutare e filtrare candidature, ecc.).

Si può dire, quindi, che è contemplato e compatibile con il regolamento l'uso di sistemi di IA in tutte le forme di lavoro, sia autonomo che subordinato, per l'espletamento di funzioni tanto nel mercato del lavoro quanto nel rapporto di lavoro ed in tutte le fasi di vita del rapporto di lavoro, da quella pre-assuntiva e selettiva, a quella del trattamento, gestione, direzione, controllo, valutazione e cessazione dei rapporti contrattuali.

Queste applicazioni in realtà non sono classificabili sempre come ad alto rischio, che costituisce il terreno elettivo e quasi esclusivo di applicazione della disciplina regolamentare europea. Nella versione finale del regolamento è stata attenuata l'originaria presunzione assoluta di alto rischio per i sistemi di IA nel campo delle risorse umane. Tale presunzione è stata degradata da assoluta a semplice e resa derogabile.

Il rischio alto, infatti, si presume soltanto se il sistema di IA presenta un rischio "significativo" di danno per la salute, la sicurezza o per i diritti fondamentali delle persone fisiche e se influenza materialmente il processo decisionale ⁽¹¹⁾. Il che comporta che, in fondo, la classificazione ad alto rischio dei sistemi di IA impiegati nel campo delle risorse umane – anche di quelli elencati nell'allegato III – non è assoluta, né automatica, ma dipende nel concreto da come è configurato il sistema, da come esso si inserisce nel processo decisionale e dalla sua rilevante rischiosità, ovvero dalla probabilità di arrecare un danno "significativo" ai diritti fondamentali delle persone ⁽¹²⁾. Circostanze, queste, che – anche in attesa che la Commissione dia corso alle esemplificazioni dei casi d'uso ai sensi dell'articolo 6, § 5 – spetta in realtà al produttore valutare e documentare prima dell'immissione sul mercato (articolo 6, § 4) e classificare

⁽¹¹⁾ Cfr. in proposito le deroghe al § 2 dell'art. 6 previste dal § 3 dello stesso articolo.

⁽¹²⁾ Ai sensi dell'art. 3, § 1, punto 2, del regolamento per "rischio" si intende «la combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso».

al momento della registrazione nell'apposita banca dati dell'UE dei sistemi di IA (articoli 49 e 71).

Ne discende che l'individuazione della classe di rischio dei sistemi di IA applicati alle risorse umane è affidata in definitiva all'autovalutazione (*self-assessment*) del produttore nella preliminare valutazione del rischio (sull'approccio *risk-based* del regolamento si veda, *infra*, § 4). Rispetto a tale classificazione il datore di lavoro o l'agenzia del lavoro o altro operatore del mercato del lavoro, nella sua veste di *deployer*, cioè utilizzatore del sistema di IA, è un soggetto passivo che prende atto della classificazione operata dal produttore. Spetta a quest'ultimo fornire al *deployer* le informazioni tecniche necessarie a rendere sufficientemente trasparente il funzionamento al fine di mettere l'utilizzatore in grado di interpretare l'output del sistema, utilizzarlo adeguatamente ed assolvere i propri obblighi (articolo 13).

In tale contesto il *deployer* prende atto delle caratteristiche tecniche e funzionali del sistema di IA così come identificate, valutate e monitorate dal produttore e si impegna ad attenersi ad esse, a monitorarne il funzionamento sulla base delle istruzioni per l'uso ricevute e, se del caso, informare il fornitore di rischi e malfunzionamenti che abbia rilevato o di elementi rilevanti per il monitoraggio della conformità (articoli 26, 72 e 79).

Al fine di inquadrare correttamente i sistemi di IA come ad alto rischio il regolamento aiuta a identificare quelle applicazioni che non sono destinate a influenzare materialmente il processo decisionale e per tale motivo non sono classificabili come ad alto rischio. Tali sono, ai sensi del § 3 dell'articolo 6, quelle applicazioni in cui l'IA è destinata a svolgere un compito marginale o non determinante ai fini decisionali in quanto si limita a: a) eseguire un compito procedurale limitato; b) migliorare il risultato di un'attività umana precedentemente completata; c) rilevare schemi decisionali o deviazioni da schemi decisionali precedenti e non è finalizzato a sostituire o influenzare la valutazione umana precedentemente completata senza un'adeguata revisione umana; d) eseguire un compito preparatorio per una valutazione pertinente.

L'esemplificazione rende evidente che la non integrale automatizzazione del sistema decisionale non è un elemento di per sé idoneo a sottrarlo dall'area di rischio alto. Può presentare un rischio non alto anche un sistema integralmente automatizzato quando si pone in posizione di complementarietà con processi decisionali umani, perché appunto deputato a

integrarsi con un sistema decisionale più ampio oppure quando sia destinato ad ottimizzare gli esiti di processi decisionali umani conclusi. Sono invece da considerarsi sempre ad alto rischio i sistemi di IA che effettuano una profilazione di persone fisiche (articolo 6, § 3). Sul piano pratico, considerati i possibili casi d'uso, come subito si vedrà, è molto probabile che un sistema di IA applicata alle risorse umane si basi su una preventiva profilazione delle persone fisiche ⁽¹³⁾ e, in tal caso, scatta la presunzione – in questo caso assoluta – di alto rischio del sistema.

3. Le applicazioni concrete della IA nel campo delle risorse umane: l'IA come strumento di lavoro e come strumento di esercizio di poteri privati

A livello applicativo i casi d'uso sono numerosi sia in Italia che all'estero e in progressiva espansione ⁽¹⁴⁾. Ci si limita in questa sede ad accennare, senza alcuna pretesa di esaustività, ad alcuni schematici esempi al solo fine di contestualizzare la disamina che si sta svolgendo e, a tale scopo, pare possibile individuare almeno quattro macro-aree applicative.

La prima area riguarda gli strumenti di IA utilizzati nei processi di pre-selezione, selezione e *matching* tra domanda e offerta e la pianificazione delle politiche attive e formative. In questo ambito sono state messe a punto diverse soluzioni per generare in automatico e ottimizzare le *job description* delle offerte di lavoro e degli annunci (*job post*), per effettuare ricerche avanzate e massive sui curriculum di possibili candidati (*CV intelligence*); strumenti per individuare e misurare i gradi di affinità e *matching score* tra candidati e *vacancy* con connessi suggerimenti per colmare i disallineamenti; strumenti di *talent scouting*; strumenti di

⁽¹³⁾ Il significato da attribuire al termine profilazione non è espressamente previsto dal regolamento sull'IA e va quindi ricavato dal regolamento (UE) 2016/679 come «qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica» (art. 4).

⁽¹⁴⁾ A. INGRAO, voce *Contratto di lavoro e digitalizzazione*, in *Enc. Dir. – I Tematici*, 2023, vol. VI, *Contratto di lavoro*, p. 245 ss.; F.V. PONTE, *Intelligenza artificiale e lavoro. Organizzazione algoritmica, profili gestionali, effetti sostitutivi*, Giappichelli, 2024.

assessment, di valutazione del potenziale e di analisi avanzata dello *skill set* (offerto/domandato) associato alle professioni; strumenti per l'analisi dei fabbisogni professionali ⁽¹⁵⁾ e competenziali anche in ottica predittiva di persone, gruppi e settori.

La seconda area di applicazione riguarda i processi di assunzione e selezione, al fine di potenziare strumenti di analisi dei candidati e ottimizzare i processi di reclutamento. In questo ambito vengono utilizzati da grandi imprese e gruppi sistemi di IA per automatizzare o semi automatizzare i colloqui e la valutazione dei candidati che permettono su larga scala di indagare non soltanto le competenze possedute, ma anche il c.d. potenziale, le attitudini e le caratteristiche personali e relazionali delle persone (*human potential intelligence*). Vengono poi applicati strumenti di IA per monitorare progressivamente e fare valutazioni di impatto degli stessi processi di *on-boarding* con il fine di affinarli, migliorarne la *compliance*, eliminare o contenere bias e pregiudizi. In questo caso i sistemi di IA hanno una sorta di fine correttivo, valutativo e migliorativo dei processi selettivi.

La terza area di applicazione riguarda la gestione del personale e l'analisi ed elaborazione dei dati delle persone anche in funzione predittiva (*people analytics*). Rientrano in questo ambito strumenti di IA per la valutazione della performance individuale e di gruppo, per la costruzione dei percorsi professionali e delle carriere, per la valorizzazione del personale e la gestione delle premialità. Si possono anche avere applicazioni mirate alla migliore configurazione delle strategie personalizzate di trattenimento e fidelizzazione del personale e di valutazione predittiva del rischio di abbandono in funzione anticipatoria delle dimissioni. Rientrano in quest'ambito gestionale anche gli strumenti più classici di ottimizzazione della gestione dei turni e della formazione e l'assortimento delle squadre o gruppi di lavoro, la programmazione delle turnazioni, così come lo smistamento dei carichi di lavoro e la razionalizzazione dei flussi di attività (*workflow algorithms*).

⁽¹⁵⁾ S. CIUCCIOVINO, *L'intermediazione alla prova dello skill mismatch*, in *LD*, 2023, n. 2, p. 309 ss.; S. CIUCCIOVINO, *Professionalità, occupazione e tecnologia nella transizione digitale*, in *Federalismi.it*, 2022, n. 9, p. 129 ss.; M. FAIOLI, *Matchmaking: la tecnologia avanzata per il mercato del lavoro*, in *LD*, 2023, n. 2, p. 333 ss.; F. LAMBERTI, *Formazione, occupabilità e certificazione delle competenze (tramite blockchain): un'alternativa alla "disoccupazione tecnologica"*, in M. BIASI (a cura di), *Diritto del lavoro e intelligenza artificiale*, Giuffrè, 2024, p. 281 ss.

Infine un quarto ambito applicativo dei sistemi di IA riguarda la salute e sicurezza sul lavoro ⁽¹⁶⁾. Qui le applicazioni sono a volte sorprendenti e spaziano dai dispositivi di protezione individuale intelligenti (elmetti e dispositivi indossabili per controllare lo stato di attenzione e altre caratteristiche psico fisiche delle persone), a strumenti e dispositivi (occhiali, monitor, guanti, simulatori, ecc.) che grazie alla realtà aumentata e al metaverso ⁽¹⁷⁾ permettono di svolgere in maggiore sicurezza e con maggiore puntualità compiti delicati o rischiosi. Allo stesso tempo si hanno utilizzi dei *data analytics* per la prevenzione della incidentalità.

In realtà l'elencazione dei potenziali usi della IA alle risorse umane contenuta nell'allegato III del regolamento non è esaustiva, in quanto è mirata specificamente ad identificare i sistemi di IA che si presumono ad alto rischio, tali da richiedere il rispetto di requisiti particolari e obblighi stringenti in capo ai produttori e altri soggetti lungo la catena del valore. A ben vedere nell'allegato III vengono contemplati essenzialmente i sistemi di IA che servono ad esercitare prerogative e poteri, sia da parte di soggetti che nel mercato del lavoro svolgono funzioni di intermediazione, promozione dell'incontro domanda offerta di lavoro, di ricerca e selezione del personale e di formazione professionale; sia da parte di soggetti (datori di lavoro e committenti) che gestiscono rapporti di lavoro subordinato, autonomo o parasubordinato.

In realtà le applicazioni dell'IA di interesse lavoristico travalicano quelle elencate nell'allegato III ai §§ 3 e 4. Sia perché l'area dell'esercizio dei poteri privati che può essere effettuata attraverso l'IA è più ampia di quella presa in considerazione nell'allegato III (si pensi, solo per fare un esempio, alle applicazioni che mirano a gestire le premialità e i sistemi retributivi oppure a indirizzare le politiche di *retention* e *attraction* del personale); sia perché le applicazioni prese in esame dall'allegato non contemplano i sistemi di IA che configurano essi stessi strumenti di lavoro.

Il sistema di IA può, infatti, configurare uno strumento di lavoro quanto interviene nello svolgimento della prestazione oppure quando ha funzioni di protezione della salute e sicurezza o prevenzione degli infortuni.

⁽¹⁶⁾ M. GIOVANNONE, *Responsabilità datoriale e prospettive regolative della sicurezza sul lavoro. Una proposta di ricomposizione*, Giappichelli, 2024, p. 161 ss.

⁽¹⁷⁾ V. MAIO, *Contratto di lavoro e metaverso*, in M. BIASI (a cura di), *op. cit.*, p. 581 ss.; F. LAMBERTI, *Il metaverso: profili giuslavoristici tra rischi nuovi e tutele tradizionali*, in *Federalismi.it*, 2023, n. 4.

I problemi che si pongono dal punto di vista giuslavoristico sono differenti a seconda che i sistemi di IA rappresentino uno strumento di esercizio di poteri e prerogative oppure uno strumento per l'esecuzione della prestazione. Nel secondo caso la persona che lavora non è soggetto passivo del sistema di IA, bensì è parte attiva, è un agente dell'IA. È chiaro che già questa configurazione del rapporto umano/macchina pone, dal punto di vista giuridico, questioni in parte differenti. Oltre a quella, non secondaria, del possesso delle competenze necessarie a rapportarsi in modo corretto e consapevole con gli strumenti di IA, si pongono problemi sul piano della responsabilità del singolo rispetto ad un utilizzo improprio o una interpretazione errata degli output algoritmici; ma anche questioni relative al particolare atteggiarsi del dovere di diligenza e collaborazione, dell'obbligo di obbedienza, della professionalità, della tutela della salute e sicurezza e della responsabilità per i danni eventualmente procurati ⁽¹⁸⁾.

Se invece guardiamo ai sistemi di IA come mezzi di esercizio di prerogative e poteri datoriali, vengono in rilievo principalmente le questioni dell'esercizio appropriato del potere e dei relativi limiti ⁽¹⁹⁾. Nonché il ripensamento che l'IA sollecita sul piano della garanzia di esercizio non

⁽¹⁸⁾ M. FAIOLI, *Mansioni e macchina intelligente*, Giappichelli, 2018; M. FAIOLI, *Data analytics, robot intelligenti e regolazione del lavoro*, in *Federalismi.it*, 2022, n. 9, p. 149; M. FAIOLI, *Prospects on Risks, Liabilities and Artificial Intelligence, empowering Robots at Workplace Level. The EU Regulation 2024/1689, with the related EU and Domestic Legal Frames, compared to the U.S. Legal System (September 01, 2024)*, Working Paper Fondazione Giacomo Brodolini, 2024, n. 24; L. ZAPPALÀ, *Informatizzazione dei processi decisionali e diritto del lavoro: algoritmi, poteri datoriali e responsabilità del prestatore nell'era dell'intelligenza artificiale*, in *Biblioteca '20 Maggio'*, 2021, n. 2.

⁽¹⁹⁾ A. BELLAVISTA, R. SANTUCCI (a cura di), *Tecnologie digitali, poteri datoriali e diritti dei lavoratori*, Giappichelli, 2022; M. NOVELLA, *Poteri del datore di lavoro nell'impresa digitale: fenomenologia e limiti*, in *LD*, 2021, n. 3-4; R. VOZA, *Lavoro autonomo e capitalismo delle piattaforme*, Cedam, 2018; M. PERUZZI, *Intelligenza artificiale e lavoro. Uno studio su poteri datoriali e tecniche di tutela*, Giappichelli, 2023; U. GARGIULO, *Intelligenza Artificiale e poteri datoriali: limiti normativi e ruolo dell'autonomia collettiva*, in *Federalismi.it*, 2023, n. 29; L. TEBANO, *Lavoro, potere direttivo e trasformazioni organizzative*, Editoriale Scientifica, 2020; M. GIOVANNONE, *Il lavoro tramite piattaforma nell'ordinamento europeo*, in M. BIASI (a cura di), *op. cit.*, p. 497 ss.

arbitrario, razionale, non discriminatorio ⁽²⁰⁾ e trasparente del potere ⁽²¹⁾, e sul piano dei presidi che occorre apprestare per garantire, nei nuovi scenari, la protezione della dignità, personalità morale, libertà, autodeterminazione, salute e sicurezza della persona.

4. L'approccio regolativo europeo basato sul rischio, l'*accountability* algoritmica e la ridotta responsabilità dei *deployers* (datori di lavoro, agenzie per il lavoro e altri soggetti del mercato del lavoro)

I sistemi di IA sono trattati dal regolamento essenzialmente come un “prodotto” di cui occorre garantire il corretto funzionamento e la qualità tecnica. Dal punto di vista giuslavoristico ciò può apparire in parte insoddisfacente, ma il regolamento sconta le basi giuridiche ai sensi delle quali è stato adottato: gli articoli 114 e 16 TFUE ⁽²²⁾, nell'ambito della strategia dell'UE per il mercato unico digitale e secondo la logica del *new legislative framework* (NLF), diretto a garantire la conformità dei prodotti immessi nel mercato UE alla legislazione applicabile attraverso le valutazioni di conformità e l'uso della marcatura CE.

⁽²⁰⁾ M. BARBERA, *Discriminazioni algoritmiche e forme di discriminazione*, in *LLI*, 2021, n. 1, I; M. BARBERA, *Principio di eguaglianza e divieti di discriminazione*, in M. BARBERA, A. GUARISO (a cura di), *La tutela antidiscriminatoria. Fonti, strumenti, interpreti*, Giappichelli, 2019, p. 59 ss.; M.V. BALLESTRERO, *Ancora sui rider. La cecità discriminatoria della piattaforma*, in *Labor*, 2021, n. 1, p. 104 ss.; C. ALESSI, *Lavoro tramite piattaforma e divieti di discriminazione nell'UE*, in C. ALESSI, M. BARBERA, L. GUAGLIANONE (a cura di), *Impresa, lavoro e non lavoro nell'economia digitale*, Cacucci, 2019; A. PERULLI, *La discriminazione algoritmica: brevi note introduttive a margine dell'ordinanza del Tribunale di Bologna*, in *LDE*, 2020, n. 1, p. 7; A. LO FARO, *Algorithmic Decision Making e gestione dei rapporti di lavoro: cosa abbiamo imparato dalle piattaforme*, in *Federalismi.it*, 2022, n. 25; G. GAUDIO, *Algorithmic management, poteri datoriali e oneri della prova: alla ricerca della verità materiale che si cela dietro l'algoritmo*, in *LLI*, 2020, n. 2, p. 21 ss.

⁽²¹⁾ G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *PD*, 2019, n. 2, p. 199 ss.

⁽²²⁾ A. ALAIMO, *Il Regolamento sull'Intelligenza Artificiale: dalla proposta della Commissione al testo approvato dal Parlamento. Ha ancora senso il pensiero pessimistico?*, in *Federalismi.it*, 2023, n. 25; M. PERUZZI, *Intelligenza artificiale e lavoro: l'impatto dell'AI Act nella ricostruzione del sistema regolativo UE di tutela*, in M. BIASI (a cura di), *op. cit.*, p. 113 ss.

La base giuridica utilizzata non è quella tipica degli interventi di politica sociale, ed anzi è escluso espressamente che l'armonizzazione legislativa adottata con lo strumento dell'articolo 114 TFUE possa essere applicabile alla materia dei «diritti ed interessi dei lavoratori dipendenti». Ciò spiega il motivo per il quale il regolamento è principalmente focalizzato sui requisiti dei sistemi di IA come prodotto e prevalentemente sugli obblighi dei produttori che li sviluppano ed immettono sul mercato.

In tale prospettiva la posizione del *deployer*, cioè il datore di lavoro o, comunque, il soggetto del mercato del lavoro che ne fa uso nel campo delle risorse umane, dell'occupazione e delle politiche attive (agenzie del lavoro, di somministrazione, di selezione, enti formativi, bilaterali e altri soggetti impegnati nelle politiche attive del lavoro), che utilizza i sistemi di IA fabbricati da altri, rimane in una posizione secondaria e funzionalmente dipendente da quella del produttore.

Il ruolo strategico del produttore, rispetto al *deployer* utilizzatore, è esaltato anche dalla tecnica regolativa *risk-based* adottata dal regolamento⁽²³⁾, già sperimentata per altri importanti materie di fonte europea di interesse giuslavoristico, come la salute e sicurezza sul lavoro e la protezione dei dati personali.

Questo approccio regolativo, incentrato su garanzie procedurali tipiche dei sistemi di *accountability* nello spazio europeo della privacy digitale, vede nel *due data process* lo strumento più efficace per garantire il rispetto delle regole a tutti i livelli e così porre argini al dilagare di poteri digitali privati, responsabilizzando i soggetti economici nelle catene di valore del mercato globale. Questa tecnica, già presente nell'apparato normativo della protezione dei dati personali, avrebbe l'ambizione anche di travalicare i limiti dei confini territoriali di applicazione delle norme prescrittive nello spirito di massima estensione territoriale possibile (*territorial extension*) delle politiche europee in questa materia.

Anche il regolamento opta per una tecnica in parte immediatamente prescrittiva ed in parte basata sulla responsabilizzazione (*accountability*) e sulla valutazione preventiva del rischio (*risk-based*) dei soggetti obbligati. Questa tecnica è funzionale a permettere la massima adattabilità delle regole a contesti aziendali molto differenti di applicazione dei

⁽²³⁾ P. LOI, *Il rischio proporzionato nella proposta di regolamento sull'IA e i suoi effetti nel rapporto di lavoro*, in *Federalismi.it*, 2023, n. 4; M. PERUZZI, *Intelligenza artificiale e lavoro: l'impatto dell'AI Act nella ricostruzione del sistema regolativo UE di tutela*, cit.

sistemi di IA e alle situazioni concrete enormemente variegata di traduzione applicativa, anche sul piano tecnico-funzionale, che renderebbe assai complesso, soprattutto in uno scenario di rapido sviluppo tecnologico, dettare a livello legislativo regole uniformi e di dettaglio immediatamente prescrittive.

Per tale motivo il regolamento abbraccia una tecnica ormai diventata consueta che si basa, da un lato, sulla preliminare individuazione di ruoli e soggetti obbligati (produttore, *deployer*, rappresentante autorizzato, importatore, distributore, operatore) con i rispettivi obblighi e responsabilità; e dall'altro lato, sulla responsabilizzazione organizzativa, mediante l'adozione di un processo gestionale circolare ed iterativo che assicura il controllo e la conformità lungo l'intero ciclo di vita e la catena del valore della IA ⁽²⁴⁾.

È pur vero che questa impostazione nella versione definitiva del regolamento è stata ridimensionata rispetto alle versioni originarie. Infatti l'approccio gestionale organizzativo vincola ora esclusivamente i produttori/fornitori e non è impegnativa per i *deployers*, ma soltanto da questi eventualmente adottata in modo volontario attraverso codici di condotta: cfr. articolo 95, § 2). Inoltre l'approccio gestionale organizzativo è limitato ai sistemi di IA ad alto rischio, e pertanto non è vincolante per quei sistemi che trovano applicazione nel campo dell'occupazione e gestione delle risorse umane che non sono classificabili ad alto rischio.

Nel caso di sistemi di IA ad alto rischio il produttore deve dotarsi di un sistema di gestione dei rischi fin dalla progettazione, pianificato ed eseguito nel corso dell'intero ciclo di vita di un sistema di IA (articolo 9); oltre che di un appropriato sistema di gestione e governance dei dati (articolo 10) e di un sistema di gestione della qualità che garantisca la conformità al regolamento attraverso un sistema documentato e ordinato di politiche, procedure e istruzioni scritte tali da garantire costantemente la complessiva *compliance* del sistema (articolo 17).

Va sottolineato che la conformità al regolamento del sistema di gestione della qualità è autodichiarato dal produttore in base ad una procedura di autovalutazione e controllo interno (effettuata ai sensi dell'allegato VI) in tutti i sistemi ad alto rischio di cui ai punti 2-8 dell'allegato III (compresi quindi quelli relativi all'occupazione e le risorse umane), mentre l'obbligo di certificazione del sistema di gestione da parte di un

⁽²⁴⁾ Per una critica, M. BARBERA, "La nave deve navigare". *Rischio e responsabilità al tempo dell'impresa digitale*, in *LLI*, 2023, n. 2.

organismo terzo notificato è richiesto unicamente per i sistemi di IA ad alto rischio che fanno uso della biometria (cfr. articolo 43).

Il sistema di gestione dei rischi richiede un riesame e un aggiornamento costanti e sistematici e adotta il classico approccio organizzativo, che parte dalla preliminare mappatura e analisi dei rischi, da quelli noti e ragionevolmente prevedibili, a quelli che possono derivare da un utilizzo improprio ragionevolmente prevedibile, a quelli messi in evidenza dal sistema di monitoraggio successivo all'immissione sul mercato, dove la partecipazione e collaborazione del *deployer* è prevista, ma non è indispensabile (cfr. articolo 72).

In sostanza dal produttore dipende sia l'individuazione dei rischi sia l'adozione delle conseguenti misure opportune di gestione, che non potranno quindi che essere personalizzate, adeguate al contesto lavorativo, concrete e sottoposte a rivalutazione periodica. Si tratta di adottare una sorta di DVR (documento di valutazione del rischio) della IA e un modello organizzativo gestionale molto simile a quello che informa il sistema della salute e sicurezza sul luogo di lavoro e con una tecnica simile a quello che presiede alla valutazione preliminare di impatto del *data protection impact assessment* (DPIA) nell'ambito del sistema di protezione dei dati personali⁽²⁵⁾. Con la non trascurabile differenza che, nei contesti normativi della sicurezza e della privacy, la responsabilizzazione incombe principalmente sul datore di lavoro in quanto direttamente investito delle leve gestionali del personale e titolare del trattamento dei dati, mentre nel contesto della IA la responsabilità principale è posta in capo al fornitore dei sistemi, anche se questi vengono utilizzati e applicati al personale dal *deployer*. Ciò è coerente con l'impostazione del regolamento che guarda al sistema di IA come prodotto e che lascia in secondo piano l'IA come processo di gestione di relazioni tra soggetti privati e di esercizio di poteri verso le risorse umane.

Sul *deployer* datore di lavoro o altro soggetto del mercato del lavoro che utilizza il sistema di IA incombono piuttosto obblighi di secondo livello,

⁽²⁵⁾ P. TULLINI, *Dati*, in M. NOVELLA, P. TULLINI (a cura di), *Lavoro digitale*, Giappichelli, 2022; A. TROJSI, *Sull'impatto giuslavoristico del Data Governance Act. Riflessioni sistemiche a prima lettura del Regolamento (UE) 2022/868*, in *Federalismi.it*, 2023, n. 4; S. CIUCCIOVINO, *Trattamento dei dati nell'ambito dei rapporti di lavoro*, in R. D'ORAZIO, G. FINOCCHIARO, O. POLLICINO, G. RESTA (a cura di), *Codice della privacy e data protection*, Giuffrè, 2021, p. 947 ss. 1; A. INGRAO, *La protezione dei dati personali dei lavoratori nel diritto vivente al tempo degli algoritmi*, in A. BELLAVISTA, R. SANTUCCI (a cura di), *op. cit.*, p. 127 ss.

derivati dal produttore, e, comunque, principalmente per i sistemi ad alto rischio, su cui ci si soffermerà sul prossimo paragrafo. Per il resto gran parte delle tutele, ad iniziare da quelle relative alla trasparenza verso le persone interessate, sono realizzate in base al regolamento (UE) 2016/679 sulla protezione dei dati personali, ove applicabile, con il quale il regolamento sulla IA deve coordinarsi ed integrarsi. Il regolamento (UE) 2024/1689 quindi non è una fonte europea esaustiva dell'impiego della IA, in quanto fa sistema con altre fonti, in primo luogo il GDPR e le norme dei Trattati e della Carta dei diritti fondamentali dell'Unione su cui si radicano ormai i diritti fondamentali a livello europeo.

Nel testo definitivo del regolamento esce in parte depotenziato quel principio di *eXplainable AI* (XAI), cioè di spiegabilità dell'IA a tutto tondo, che nelle ambizioni iniziali doveva costituire un principio conduttore della regolazione europea per una IA affidabile e sicura e, in realtà, oggi risulta relegato essenzialmente nell'ambito dei sistemi ad alto rischio o per finalità generali con rischio sistemico (vi veda *infra*, §§ 5 e 6).

5. Legalità ed etica algoritmica nel rapporto di lavoro ai sensi del quadro regolativo europeo sull'IA

Il dibattito e gli atti preparatori che hanno preceduto e accompagnato il varo del regolamento (UE) 2024/1689 in merito all'etica e ai valori che dovrebbero informare lo sviluppo dell'IA sono molto ricchi. Certamente si è avuto uno sviluppo impetuoso della sensibilità delle istituzioni e del dibattito pubblico verso i valori dell'antropocentrismo e del rispetto della persona umana, dell'autonoma determinazione, della libertà e dignità della persona al cospetto del rischio di una disumanizzazione indotta dalla crescente esposizione alla IA.

A livello europeo si è approdati alla Dichiarazione europea sui diritti e i principi digitali per il decennio digitale proclamata il 23 gennaio 2023 dal Parlamento europeo, dal Consiglio e dalla Commissione e alla emanazione degli *Orientamenti etici per un'IA affidabile* elaborato dal gruppo di esperti ad alto livello sull'IA (AI HLEG) che sono sfociati poi nella comunicazione della Commissione *Creare fiducia nell'intelligenza artificiale antropocentrica* dell'8 aprile 2019 (COM(2019)168 final); orientamenti etici richiamati in più punti dal regolamento.

A livello ancora più generale, occorre ricordare poi la citata *Council of Europe Framework Convention* ⁽²⁶⁾, che stabilisce i principi che dovranno essere osservati dai Paesi sottoscrittori lungo l'intero ciclo di vita dei sistemi di IA nel rispetto dei principi di: dignità umana e autonomia individuale; trasparenza e sorveglianza umana; accountability e responsabilità; uguaglianza e non discriminazione; privacy e protezione dei dati; affidabilità e sicurezza nell'innovazione.

A livello UE in particolare l'AI HLEG ha individuato sette principi etici per una IA affidabile, etica ed antropocentrica, richiamati dal considerando 27 del regolamento UE che, però, ne sottolinea contestualmente anche la non vincolatività. I sette principi etici sono così elencati: intervento e sorveglianza umani; robustezza tecnica e sicurezza; vita privata e governance dei dati; trasparenza; diversità, non discriminazione ed equità; benessere sociale e ambientale; responsabilità. Tali principi in verità toccano diritti fondamentali affermati dalla Carta dei diritti fondamentali e valori essenziali su cui si fonda l'Unione, sicché appare riduttivo qualificarli soltanto come principi etici. In larga misura rinviano, implicitamente o esplicitamente, ad ulteriori diritti e valori fondamentali dell'UE, come ad esempio il rispetto della dignità umana e dell'autonomia personale, sotteso, secondo quanto previsto dal considerando 27 del regolamento, al criterio della sorveglianza umana.

Pur volendo mantenerli ascritti al terreno dei valori etici, senza farli assurgere al rango di regole giuridiche da rispettare, essi sono comunque da osservare almeno come canoni interpretativi delle norme europee e del regolamento pertinenti in materia di IA. Inoltre per la loro stretta connessione con i diritti fondamentali sanciti dalla Carta concorrono a definire i vincoli posti direttamente dai diritti fondamentali dell'UE allo sviluppo e applicazione dei sistemi di IA (cfr. articolo 1, § 1, regolamento). In aggiunta a ciò i diritti fondamentali rilevano nella definizione dello stesso ambito di applicazione del regolamento, atteso che il "rischio" necessario ad identificare e classificare i sistemi di IA ai sensi del regolamento ⁽²⁷⁾ dipende, come si è detto, dal «rischio *significativo* di danno

⁽²⁶⁾ La Convenzione ha un campo di applicazione ben più ampio dell'Unione europea ed è aperta all'adesione anche di Stati non aderenti al Consiglio d'Europa. Hanno aderito alla data di redazione del presente scritto anche Stati Uniti, Regno Unito, Israele, Georgia, Islanda, Andorra, Norvegia, Moldavia, San Marino.

⁽²⁷⁾ L'art. 79, § 1, del regolamento prevede che «un sistema di IA che presenta un rischio è inteso come un "prodotto che presenta un rischio" quale definito all'articolo 3, punto 19, del regolamento (UE) 2019/1020 nella misura in cui presenta rischi per la

per la salute, la sicurezza o i diritti fondamentali delle persone fisiche». Il criterio dell'alto rischio rappresenta in definitiva il filtro applicativo di molta parte delle prescrizioni del regolamento.

Ecco allora che principi etici, diritti fondamentali e obblighi giuridici posti dal regolamento (e dalle altre fonti esterne ad esso, con questo integrate) si intrecciano a tessere la trama normativa e gli argini entro i quali l'innovazione tecnologica potrà svilupparsi in Europa, nella ricerca dell'equilibrio mobile tra innovazione, libertà di impresa e tutele fondamentali.

Dal punto di vista strettamente tecnico giuridico i criteri europei di legalità della IA sono quelli previsti dalla parte dispositiva del regolamento, alle condizioni e nei limiti ivi disciplinati.

Cionondimeno i criteri etici, secondo quanto previsto dallo stesso regolamento, dovrebbero trovare applicazione, ove possibile, nella progettazione dei sistemi di IA e dovrebbero ispirare i codici di condotta e lo sviluppo da parte dei diversi portatori di interesse delle migliori pratiche e norme volontarie di attuazione del regolamento (cfr. considerando 27). Non a caso il riferimento ai principi etici compare nell'articolo 95 del regolamento al § 2 nell'ambito degli obiettivi e criteri di prestazione da prevedere nei codici di condotta per l'applicazione volontaria da parte dei *deployers* di requisiti dei sistemi di IA (sul punto si tornerà *infra*, § 7).

È peculiare che soltanto in tale contesto volontario compaia il riferimento esplicito al criterio della «valutazione e la prevenzione dell'impatto negativo dei sistemi di IA sulle persone vulnerabili o sui gruppi di persone vulnerabili, anche per quanto riguarda l'accessibilità per le persone con disabilità, nonché sulla parità di genere». Il che significa che la valutazione di impatto e la prevenzione degli impatti negativi discriminatori non rappresentano in senso tecnico un criterio di legalità dei sistemi di

salute o la sicurezza o per i diritti fondamentali delle persone». L'art. 3, § 19, del regolamento (UE) 2019/1020 definisce "prodotto che presenta un rischio" «un prodotto che potenzialmente potrebbe pregiudicare la salute e la sicurezza delle persone in generale, la salute e la sicurezza sul posto di lavoro, la protezione dei consumatori, l'ambiente e la sicurezza pubblica, nonché altri interessi pubblici tutelati dalla normativa di armonizzazione dell'Unione applicabile». A sua volta l'art. 6, § 3, del regolamento IA esenta dalla classificazione come ad alto rischio un sistema di IA compreso nell'allegato III «se non presenta un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche».

IA, bensì un principio etico non strettamente vincolante, da osservare nel caso di codici volontari di condotta da parte del *deployer*.

È pur vero che l'articolo 1 del regolamento nel definire il suo scopo afferma di voler garantire una IA «antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, contro gli effetti nocivi dei sistemi di IA nell'Unione», ma al contempo il regolamento contiene al proprio interno poche norme vincolanti direttamente preordinate a questo scopo e piuttosto rinvia alle altre disposizioni di diritto dell'Unione che restano impregiudicate e con le quali il regolamento va dunque coordinato (si veda ad esempio l'articolo 5, § 8, sui divieti; l'articolo 50, § 6, sugli obblighi di trasparenza per i *deployers*). In aggiunta a ciò il regolamento rinvia alle condizioni più favorevoli ai lavoratori in termini di tutela dei loro diritti in relazione all'uso dei sistemi di IA da parte dei datori di lavoro che l'UE o i singoli Stati membri possono adottare per via legislativa, amministrativa o contrattuale collettiva (articolo 2, § 11).

Può dirsi quindi che il quadro delle tutele sociali nel contesto della IA si compie grazie alla complementarità del regolamento (UE) 2024/1689 rispetto ai sistemi nazionali di protezione sociale e grazie alla interdipendenza del regolamento con la disciplina privacy⁽²⁸⁾.

Volendo evidenziare gli obblighi che il regolamento pone direttamente in capo al datore di lavoro o altro soggetto del mercato del lavoro (agenzie del lavoro, agenzie formative, ecc.), nella loro veste di *deployer* di sistemi di IA da applicare alle risorse umane, occorre distinguere gli obblighi a valenza generale, quelli che riguardano i sistemi ad alto rischio e quelli che riguardano i sistemi destinati ad interagire con le persone indipendentemente dalla classe di rischio.

Gli obblighi a valenza generale appaiono abbastanza circoscritti anche se molto rilevanti e consistono nell'alfabetizzazione e formazione verso

(28) In realtà la relazione di interdipendenza tra discipline della IA e disciplina della protezione dei dati personali è tutt'altro che semplice. Ad iniziare dalle competenze rivendicate dalle rispettive authority. Basti guardare alla posizione del nostro Garante per la protezione dei dati personali che rivendica a sé un'ampia competenza in materia di IA. Sul punto si veda F. LAMBERTI, *La proposta di regolamento UE sull'Intelligenza Artificiale alla prova della privacy*, in www.federalismi.it, 29 giugno 2022, e più recentemente il parere 2 agosto 2024 reso dal Garante sullo schema di disegno di legge recante disposizioni e deleghe in materia di intelligenza artificiale.

il personale (articolo 4). La previsione è significativamente inserita nel contesto delle disposizioni generali e obbliga i *deployers* a garantire un livello “sufficiente” di alfabetizzazione e formazione del loro personale «nella misura del possibile» anche verso qualsiasi altra persona che si occupa del funzionamento e dell’utilizzo dei sistemi di IA per loro conto, «prendendo in considerazione le loro conoscenze tecniche, la loro esperienza, istruzione e formazione, nonché il contesto in cui i sistemi di IA devono essere utilizzati, e tenendo conto delle persone o dei gruppi di persone su cui i sistemi di IA devono essere utilizzati». È possibile dedurre da tale previsione un obbligo vincolante di fonte europea in capo ai *deployers* di tutti i sistemi di IA di farsi carico della formazione personalizzata che sembra orientato a garantire oltre la consapevolezza, il corretto funzionamento dei sistemi. Non a caso l’obbligo formativo è esteso ai soggetti incaricati dal datore di lavoro del funzionamento dei sistemi. La previsione, rientrando nel capo I del regolamento, entra in vigore a decorrere dal 2 febbraio 2025 (articolo 113).

Gli obblighi connessi esclusivamente ai sistemi ad alto rischio riguardano invece: la preventiva informazione sindacale (articolo 26, § 7) ⁽²⁹⁾ e l’informativa individuale (articolo 26, § 11) ⁽³⁰⁾; il controllo sui dati nel caso il *deployer* eserciti il controllo sui dati di input (articolo 26, § 4) per garantirne la pertinenza e la sufficiente rappresentatività ⁽³¹⁾; la sorveglianza umana (articolo 14); la spiegabilità dei processi decisionali (articolo 86).

⁽²⁹⁾ Nei sistemi di IA ad alto rischio l’obbligo di preventiva informazione sindacale richiede che «prima di mettere in servizio o utilizzare un sistema di IA ad alto rischio sul luogo di lavoro, i *deployer* che sono datori di lavoro informano i rappresentanti dei lavoratori e i lavoratori interessati che saranno soggetti all’uso del sistema di IA ad alto rischio. Tali informazioni sono fornite, se del caso, conformemente alle norme e alle procedure stabilite dal diritto e dalle prassi dell’Unione e nazionali in materia di informazione dei lavoratori e dei loro rappresentanti».

⁽³⁰⁾ A livello individuale, verso i lavoratori si applica un obbligo di informazione in quanto «i *deployer* dei sistemi di IA ad alto rischio di cui all’allegato III che adottano decisioni o assistono nell’adozione di decisioni che riguardano persone fisiche informano queste ultime che sono soggette all’uso del sistema di IA ad alto rischio».

⁽³¹⁾ Si tratta di un adempimento importante e strategico, anche al fine di prevenire fenomeni di discriminazione algoritmica e storica. Ma, come detto, ad applicazione limitata ai sistemi ad alto rischio. Sul tema in generale si veda M. PERUZZI, *Il diritto antidiscriminatorio al test di intelligenza artificiale*, in *LLI*, 2021, n. 1, I.; M. PERUZZI, *Intelligenza artificiale e lavoro. Uno studio su poteri datoriali e tecniche di tutela*, cit.

Per quanto riguarda infine gli obblighi per i particolari sistemi destinati a interagire direttamente con le persone fisiche, non limitati a quelli ad alto rischio, sono previsti essenzialmente obblighi di trasparenza (articolo 50). In questo caso si tratta di una specifica accezione di trasparenza di interazione che, come si vedrà in seguito, costituisce in realtà soltanto una delle possibili dimensioni della trasparenza algoritmica.

Da questa rapida panoramica si capisce anche come principi rilevanti per il c.d. antropocentrismo, connessi direttamente ai valori etici enunciati dagli Orientamenti dell'AI HLEG, cioè la trasparenza e la spiegabilità, e richiamati anche come principi essenziali nella Convenzione del Consiglio d'Europa, in realtà nel regolamento assumono valore di regola vincolante soltanto per i sistemi ad alto rischio e per i particolari sistemi destinati ad interagire con le persone fisiche.

In conclusione la prospettiva dei destinatari finali del processo decisionale mediato dalla IA – lavoratori e lavoratrici – è soddisfatta in parte dal regolamento e in larga parte dalle altre fonti vincolanti applicabili: la Carta dei diritti fondamentali, i Trattati e altri regolamenti UE, il GDPR, da leggere e interpretare alla luce del corposo decalogo di principi etici che permeano l'inarrestabile ascesa dell'antropocentrismo della IA. Così pure un ruolo importante in questa direzione è svolto dall'apparato di governance, vigilanza e controllo a livello nazionale e europeo apprestato dal regolamento (cfr. capo VII e VIII).

Chiaro è che, in questo contesto, il terreno è fertile per lo sviluppo di quella tendenza giurisprudenziale fondata sui principi e diritti fondamentali (c.d. rimedialismo giurisprudenziale) ormai da tempo diffusa, che amplia notevolmente gli spazi del sindacato giudiziale costituzionale diffuso⁽³²⁾.

Ma sarebbe un errore pensare di affidare al contenzioso e all'interpretazione giurisprudenziale la delimitazione dei limiti sociali allo sviluppo della IA. Così come sarebbe errato pensare di definire in regole prescrittive di dettaglio un fenomeno in così rapida evoluzione, mentre occorre guardare a nuove tecniche che passano per un diverso sistema di generazione di regole e di limiti, in larga parte affidato al controllo preventivo e permanente da parte di authority, organismi di vigilanza, condotte volontarie oltre che sistemi di verifica e analisi di impatto che

⁽³²⁾ Sul tema si permetta il rinvio a S. CIUCCIOVINO, *La crisi della fattispecie e l'ap-proccio rimediale nella discussione giuslavoristica*, in *DLM*, 2024, n. 1, p. 5 ss.

contribuiscono a tratteggiare un nuovo sistema di controllo della libertà economica privata negli scenari globali.

6. Il diritto all'informazione e alla trasparenza come principio trasversale a bilanciamento del *potere tecnologico*: finzioni e funzioni della trasparenza algoritmica

Non c'è dubbio che tra i principali presidi dei diritti delle persone nei confronti della IA e del potere digitale ⁽³³⁾, un posto di primo piano è occupato dalla trasparenza. Obblighi di trasparenza, sub specie di diritti di informazione e diritti di accesso, *ex ante* ed *ex post*, sono previsti oltre che nel regolamento IA, nell'ambito del GDPR, della direttiva (UE) 2019/1152 del 20 giugno 2019, relativa a condizioni di lavoro trasparenti e prevedibili nell'Unione europea e nella proposta di direttiva relativa al miglioramento delle condizioni di lavoro nel lavoro mediante piattaforme digitali (COM(2021)762 final) ⁽³⁴⁾. Diverse norme sulla trasparenza sono disseminate nella legislazione interna ⁽³⁵⁾, tra le quali spicca l'articolo 1-*bis* del decreto legislativo n. 152/1997 ⁽³⁶⁾.

Appare necessario, tuttavia, operare una distinzione tra le diverse funzioni e le differenti accezioni di trasparenza.

In linea generale, con riferimento ai sistemi di IA, sembra possibile distinguere almeno tre accezioni di trasparenza: a) la trasparenza *di interazione*, che riguarda il diritto delle persone umane di sapere se, come e quando stanno interagendo con una macchina; b) la trasparenza *di*

⁽³³⁾ O. POLLICINO, voce *Potere digitale*, in *Enc. Dir. – I Tematici*, 2023, vol. V, *Potere e Costituzione*, p. 410 ss.

⁽³⁴⁾ P. TULLINI, *La Direttiva Piattaforme e i diritti del lavoro digitale*, in *LLI*, 2022, n. 1, R.

⁽³⁵⁾ A. ZILLI, *La trasparenza nel lavoro subordinato. Principi e tecniche di tutela*, Pacini, 2022.

⁽³⁶⁾ M. FAIOLI, *Trasparenza e monitoraggio digitale. Perché abbiamo smesso di capire la norma sociale europea*, in *Federalismi.it*, 2022, n. 25; C. FALERI, *Management algoritmico e asimmetrie informative di ultima generazione*, *ivi*, 2024, n. 3; R. RAINONE, *Obblighi informativi e trasparenza nel lavoro mediante piattaforme digitali*, *ivi*; A. ZILLI, *Condizioni di lavoro (finalmente) «trasparenti e prevedibili»*, in *Labor*, 2022, n. 6, p. 881 ss.; G. PROIA, *Le novità, sistemiche e di dettaglio, del decreto trasparenza*, in *MGL*, 2022, n. 3, p. 571 ss.; E. DAGNINO, *Il diritto interno: i sistemi decisionali e di monitoraggio (integralmente) automatizzati tra trasparenza e coinvolgimento*, in M. BIASI (a cura di), *op. cit.*, p. 147 ss.

funzionamento relativa alla prevedibilità ed esplicabilità dei meccanismi di all'output del sistema; c) la trasparenza *di impatto*, che implica una valutazione degli effetti dell'impiego della IA, sia a livello preventivo mediante la valutazione e ponderazione preventiva del rischio e la responsabilizzazione del titolare del trattamento, sia *ex post*, in esito all'applicazione del sistema per verificarne appunto gli impatti a livello individuale o di gruppo.

Queste forme di trasparenza scandiscono anche le diverse fasi che accompagnano l'applicazione della IA nei contesti di lavoro, dalla immissione ed elaborazione di dati di input, alla operatività degli algoritmi e dei calcoli per l'estrazione di output e decisioni, per giungere alla trasposizione di questi in concreti atti decisionali idonei a produrre effetti giuridici sui singoli.

E così la trasparenza di interazione riguarda principalmente la fase iniziale di contatto in cui si rende edotta la persona dell'utilizzo del sistema di IA; mentre la trasparenza di funzionamento comprende sia la trasparenza dei dati (con annessi adempimenti della normativa privacy e cybersecurity), sia il funzionamento dell'algoritmo (la spiegabilità del processo decisionale), sia infine l'intelligibilità e razionalità delle decisioni assunte in esito a tale processo che producono effetti sui singoli (spiegabilità dell'atto decisionale individuale). La trasparenza di impatto, infine, si pone principalmente, anche se non esclusivamente, a valle perché guarda agli effetti e chiama in causa i processi organizzativi e gestionali aziendali di monitoraggio degli esiti.

Le tecniche e gli strumenti con i quali si può rispondere alle diverse istanze di trasparenza sono altrettanto variegate.

La trasparenza di interazione poggia principalmente sulla preventiva informazione dell'utilizzo, nonché sulla etichettatura dei prodotti artificiali e dei *deep fakes*.

La trasparenza di funzionamento si affida alla informativa preventiva, in itinere ed *ex post* sui dati, parametri, metriche, logiche di programmazione e addestramento, sul diritto alla spiegazione che permette l'intelligibilità della decisione, sul diritto alla sorveglianza umana e revisione umana, sulla informazione relativa alle misure di controllo, correzione, accuratezza, robustezza e cybersicurezza e relative metriche.

La trasparenza di impatto invece richiede una capacità di monitoraggio dei dati e delle decisioni, l'adozione di approcci gestionali/organizzativi *risk-based*; la valutazione delle misure per contenere i rischi; la compilazione di registri d'uso; l'individuazione di figure responsabili del

sistema di gestione della qualità. Comprende poi diritti di informazione, a livello individuale e collettivo, sugli impatti potenzialmente o effettivamente discriminatori derivanti dalla applicazione dei sistemi di IA, unitamente alla capacità di interpretare queste informazioni.

C'è allora da chiedersi a quale delle possibili declinazioni del concetto di trasparenza si riferiscano la normativa europea e quella italiana e a quali strumenti e tecniche concrete queste fonti affidino la sua realizzazione.

Il regolamento abbraccia una specifica accezione di *trasparenza di interazione*. L'articolo 50 prevede per i particolari sistemi di IA destinati ad interagire direttamente con le persone fisiche, non limitati a quelli ad alto rischio, l'obbligo di rendere edotta la persona già dalla prima esposizione in merito al fatto di stare interagendo con un sistema di IA, a meno che ciò non risulti evidente dal punto di vista di una persona fisica ragionevolmente informata, attenta e avveduta, tenendo conto delle circostanze e del contesto di utilizzo. Si tratta di una trasparenza di contatto.

Inoltre i *deployers* di un sistema di IA che genera o manipola immagini o contenuti audio o video che costituiscono un "*deep fake*" rendono noto che il contenuto è stato generato o manipolato artificialmente. Questa forma di trasparenza (il *deep fake label* reso obbligatorio dall'articolo 50) aiuta l'utente a discriminare tra ciò che è reale e ciò che è prodotto artificialmente dalla IA.

In sostanza sembra che il regolamento si focalizzi principalmente su quella accezione di trasparenza, che corrisponde ad una esigenza di consapevolezza (o coscienza) dell'umano nel suo rapporto con la macchina intelligente, che poi chiama in causa i principi della dignità e dell'autonomia individuale.

Se invece passiamo a guardare alla diversa accezione della *trasparenza di funzionamento* della IA nei confronti dei destinatari finali, lavoratori e lavoratrici, va detto che il regolamento di per sé non pone obblighi specifici in capo al *deployer* datore di lavoro verso il personale (salvo l'obbligo di alfabetizzazione e formazione), piuttosto su questo versante rimanda alle legislazioni nazionali e alle altre fonti europee applicabili, prima fra tutte il regolamento (UE) 2016/679 sulla protezione dei dati personali. Ed infatti l'articolo 50, § 6, afferma che rimangono «imprejudicati i requisiti e gli obblighi di cui al capo III, così come gli altri obblighi di trasparenza stabiliti dal diritto dell'Unione o nazionale per i *deployer* dei sistemi di IA». Occorre allora ricavare gli eventuali obblighi di trasparenza verso la forza lavoro dagli articoli 13, 15, 22 e 88 GDPR

per il trattamento di dati personali, in quanto pertinenti, laddove i dati personali siano funzionali ai sistemi di IA utilizzati in contesto lavorativo.

Da una diversa prospettiva gli obblighi di trasparenza di funzionamento vincolano piuttosto il produttore verso il *deployer* nei sistemi ad alto rischio ex articolo 13, regolamento. Infatti nelle “istruzioni d’uso” che devono essere fornite dal produttore all’utente occorre rispettare un certo grado di trasparenza ed informare quest’ultimo riguardo le caratteristiche, le capacità e i limiti delle prestazioni del sistema di IA ad alto rischio. Si tratta di un’accezione *tecnica* della trasparenza, rilevante nei rapporti interni tra produttore e *deployer*.

D’altro canto è anche vero che l’accezione di trasparenza fatta propria dai principi etici elaborati dall’AI HLEG ha una valenza tutto sommato circoscritta e concepita in relazione alla finalità di consentire «un’adeguata tracciabilità e spiegabilità, rendendo gli esseri umani consapevoli del fatto di comunicare o interagire con un sistema di IA e informando debitamente i *deployer* delle capacità e dei limiti di tale sistema di IA e le persone interessate dei loro diritti» (cfr. considerando 27). Si tratta quindi di una trasparenza di interazione e contatto.

Così pure nella Dichiarazione europea sui diritti e i principi digitali per il decennio digitale (2023/C 23/01) con la quale il Parlamento europeo, il Consiglio e la Commissione hanno proclamato solennemente la dichiarazione comune sui diritti e i principi digitali per il decennio digitale assumendo specifici impegni, nel paragrafo *Condizioni di lavoro giuste ed eque*, si leggono precisi impegni in ordine alla garanzia che l’utilizzo della IA sul luogo di lavoro sia trasparente e segua un approccio basato sul rischio e che siano adottate le corrispondenti misure di prevenzione per mantenere un ambiente di lavoro sicuro e sano. Ma anche in questo contesto si specifica in particolare che venga «garantita la sorveglianza umana nelle decisioni importanti che interessano i lavoratori e questi ultimi siano, in generale, informati del fatto di stare interagendo con sistemi di intelligenza artificiale».

Paradossalmente più ampia è l’accezione di trasparenza rinvenibile nella citata comunicazione della Commissione, *Creare fiducia nell’intelligenza artificiale antropocentrica*, dove assume a requisito chiave nei sistemi di IA affidabili in una dimensione che guarda maggiormente ⁽³⁷⁾

⁽³⁷⁾ Il punto IV del § 2.2, intitolato *Trasparenza*, recita: «La *tracciabilità* dei sistemi di IA dovrebbe essere garantita: è importante registrare e documentare sia le decisioni

alla connessione con la spiegabilità verso le persone destinatarie dei processi decisionali.

Si comprende allora come le accezioni di trasparenza nel regolamento siano in realtà molteplici.

La trasparenza dei requisiti di funzionamento è posta in capo principalmente al produttore e trattata alla stregua di informazioni da integrare nella documentazione tecnica del prodotto (cfr. allegato IV⁽³⁸⁾ e articolo 11 del regolamento (UE) 2024/1689) nei sistemi ad alto rischio. Il *deployer* è il destinatario e non l'agente di questa trasparenza. A sua volta il *deployer* dovrà basarsi su queste informazioni fornite dal produttore per rispondere ad eventuali richieste da parte della persona fisica interessata dalla decisione adottata sulla base dell'output di un sistema di IA che si ritenga pregiudicato nei suoi diritti fondamentali (cfr. articolo 86).

adottate dai sistemi sia l'intero processo che ha prodotto le decisioni, comprese una descrizione della raccolta e dell'etichettatura dei dati e una descrizione dell'algoritmo utilizzato. In questo contesto dovrebbe essere prevista, per quanto possibile, la *spiegabilità* del processo decisionale degli algoritmi, adattata alle persone coinvolte. È opportuno proseguire la ricerca in corso volta a sviluppare meccanismi di spiegabilità. Dovrebbero anche essere disponibili spiegazioni sulla misura in cui un sistema di IA influenza e definisce il processo decisionale organizzativo, le scelte di progettazione del sistema e la logica alla base della sua diffusione, in modo da garantire la trasparenza non solo dei dati e dei sistemi, ma anche dei modelli di business».

(³⁸) L'allegato IV prevede che la documentazione tecnica dia conto degli elementi del sistema di IA e del processo di sviluppo tra cui figurano: la logica generale del sistema di IA e degli algoritmi; le principali scelte di progettazione, comprese le motivazioni e le ipotesi formulate, anche per quanto riguarda le persone o i gruppi di persone sui quali il sistema è destinato a essere utilizzato; le principali scelte di classificazione; gli aspetti che il sistema è progettato per ottimizzare e la pertinenza dei diversi parametri; la descrizione dell'output atteso e della qualità dell'output del sistema; le metodologie e le tecniche di addestramento e i set di dati di addestramento utilizzati; le modalità di ottenimento e di selezione dei dati; le procedure di etichettatura, ad esempio per l'apprendimento supervisionato e le metodologie di pulizia dei dati; le informazioni sui dati di convalida e di prova utilizzati e sulle loro principali caratteristiche. Cui si aggiungono informazioni dettagliate sul monitoraggio, sul funzionamento e sul controllo del sistema di IA, in particolare per quanto riguarda: le sue capacità e limitazioni in termini di prestazioni, compresi i gradi di accuratezza relativi a determinate persone o determinati gruppi di persone sui quali il sistema è destinato a essere utilizzato e il livello di accuratezza complessivo atteso in relazione alla finalità prevista del sistema; i prevedibili risultati indesiderati e fonti di rischio per la salute, la sicurezza e i diritti fondamentali, nonché di rischio di discriminazione in considerazione della finalità prevista del sistema di IA; le misure di sorveglianza umana necessarie in conformità dell'art. 14, comprese le misure tecniche poste in essere per facilitare l'interpretazione degli output dei sistemi di IA da parte dei *deployers*.

La trasparenza sotto forma di spiegabilità è, a ben vedere, concepita in funzione rimediale di un diritto che si assume leso. Lo dimostra il fatto che la sua previsione è inserita nella sezione IV dedicata ai mezzi di ricorso. L'articolo 86 riconosce il diritto a qualsiasi persona interessata ad «ottenere dal *deployer* spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata», quando tale persona è stata oggetto di una decisione adottata dal *deployer* sulla base dell'output di un sistema di IA ad alto rischio elencato nell'allegato III (quindi il lavoratore o la lavoratrice interessati) e che produca effetti giuridici o incida significativamente su tale persona in un modo che essa ritenga avere un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi diritti fondamentali.

Si specifica però anche la natura sussidiaria e cedevole di tale previsione, poiché tale diritto si applica soltanto se un analogo diritto non sia altrimenti previsto dal diritto dell'Unione (articolo 86, § 3). Da ciò si comprende come la spiegabilità dell'algoritmo nei sistemi di IA verso i destinatari lavoratori e lavoratrici non è un criterio legale a valenza generale, non è dovuto in modo sistematico e preventivo, ma è dovuto verso i singoli che ne facciano richiesta e che si ritengono pregiudicati nei loro diritti fondamentali, appunto come mezzo di ricorso.

La trasparenza di impatto infine è spostata sul produttore anziché sul *deployer*.

La valutazione di impatto di norma non è richiesta ai *deployers* datori di lavoro o agenti del mercato del lavoro, in quanto l'obbligo della valutazione di impatto sui diritti fondamentali è dovuta solo da parte dei soggetti pubblici per i sistemi ad alto rischio relativi a infrastrutture critiche e a servizi essenziali (articolo 27). Le informazioni ricevute dai produttori ai sensi dell'articolo 13 sono piuttosto utilizzate dai *deployers* di sistemi di IA ad alto rischio per adempiere al loro obbligo di effettuare una valutazione d'impatto sulla *protezione dei dati* personali a norma dell'articolo 35 del regolamento (UE) 2016/679 (articolo 26, § 9), se ne ricorrono i presupposti ed inoltre il regolamento IA lascia impregiudicato l'articolo 9 del regolamento (UE) 2016/679 su trattamento dei dati biometrici (articolo 26, § 10).

Riassumendo, per il regolamento, gli obblighi relativi alla trasparenza gravano principalmente sui fabbricanti verso gli importatori e i *deployers* in una accezione della trasparenza che riguarda i requisiti di *funzionamento*.

Sui *deployers* incombono obblighi di trasparenza verso le rappresentanze sindacali e i singoli lavoratori, in una accezione di trasparenza di utilizzo o di *interazione* e, soltanto nel caso di rilevata lesione dei diritti individuali, una *trasparenza di funzionamento* (spiegabilità) come mezzo di tutela dei diritti lesi. Non deve dimenticarsi, però, come più volte osservato, che in questo ambito soccorre la trasparenza dell'utilizzo dei dati personali ai sensi del GDPR, sia di carattere generale sia specificamente attinenti ai sistemi di profilazione e decisionali automatizzati che sottopongono l'interessato a una «decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona» (cfr. articolo 22, ma anche articoli 13, 14 e 15, regolamento (UE) 2016/679).

7. Dalla *hard law* alla *soft law* come tecnica di controllo degli impatti sociali dello sviluppo tecnologico: codici di condotta, autoregolamentazione e partecipazione sindacale

Il regolamento incoraggia l'elaborazione e lo sviluppo di pratiche volontaristiche per la promozione del rispetto spontaneo delle norme da parte di soggetti o per tipologie di IA che non sarebbero sottoposti alla loro necessaria applicazione. In tal modo si intende favorire, da un lato, da parte dei soggetti obbligati la conformazione al regolamento in via anticipata rispetto alle scadenze temporali previste per la sua piena operatività; dall'altro lato una espansione dell'ambito applicativo effettivo del regolamento, tanto sul piano oggettivo, verso sistemi di IA non ad alto rischio, quanto sul piano soggettivo, verso i *deployers*, che non sarebbero direttamente tenuti all'osservanza di quelle prescrizioni e di quegli obblighi che il regolamento pone in capo ai produttori.

In tale prospettiva gli Stati membri dovrebbero incoraggiare e agevolare l'elaborazione volontaria, da parte dei *deployers*, di codici di condotta come strumento di applicazione volontaria delle norme del regolamento, anche tenendo conto delle migliori pratiche del settore (articoli 95 e 50, § 7).

I codici di condotta, pur configurando fonti autonome, dovrebbero ubbidire, secondo quanto previsto dall'articolo 95 e il considerando 165, a criteri ben precisi volti a garantirne efficacia e concretezza. Dovrebbero basarsi su «obiettivi chiari e indicatori chiave di prestazione volti a

misurare il conseguimento di tali obiettivi», quindi ispirarsi alla cultura della pianificazione e del monitoraggio in base a criteri osservabili e documentabili. I codici di condotta, come esemplifica il regolamento, possono portare all'applicazione di singoli aspetti previsti negli orientamenti etici dell'UE per una IA affidabile, anzi possono rappresentare uno strumento ottimale per favorire la più ampia adozione nella UE di un'IA etica e affidabile (cfr. considerando 165 e 27).

Tra gli aspetti su cui i codici di condotta potrebbero focalizzarsi sono menzionati espressamente la valutazione e la riduzione al minimo dell'impatto dei sistemi di IA sulla sostenibilità ambientale; la promozione dell'alfabetizzazione in materia di IA, in particolare quella delle persone che si occupano dello sviluppo, del funzionamento e dell'uso dell'IA; la facilitazione di una progettazione inclusiva e diversificata dei sistemi di IA, anche attraverso la creazione di gruppi di sviluppo inclusivi e diversificati e la promozione della partecipazione dei portatori di interessi a tale processo; la valutazione e la prevenzione dell'impatto negativo dei sistemi di IA sulle persone vulnerabili o sui gruppi di persone vulnerabili, anche per quanto riguarda l'accessibilità per le persone con disabilità, nonché sulla parità di genere (art 95).

Da questo elenco si coglie l'importanza strategica che possono giocare i codici di condotta nell'ampliamento della sfera di copertura della regolazione europea e nel contenimento degli impatti sociali connessi all'evoluzione tecnologia in ambito IA nel campo delle risorse umane, oltre la logica prescrittiva delle norme legali.

I codici di condotta possono essere elaborati da singoli fornitori oppure dai *deployers* di sistemi di IA o, ancora, da organizzazioni che li rappresentano o da entrambi, anche con la partecipazione di qualsiasi portatore di interessi e delle sue organizzazioni rappresentative, comprese le organizzazioni della società civile e il mondo accademico.

Codici di condotta potrebbero quindi essere adottati su impulso delle organizzazioni rappresentative datoriali o anche in base ad accordi con le rappresentanze sindacali o, ancora, a livello di organismi paritetici e bilaterali. Si tratterebbe di una forma molto concreta e costruttiva di partecipazione sindacale, in un momento storico in cui la partecipazione appare riprendere vigore e sembra connotare una nuova stagione delle relazioni industriali.

Codici di condotta potrebbero essere adottati a livello di settore sul piano nazionale o territoriale ed essere aperti all'adesione delle singole imprese

o soggetti interessati. Oppure essere adottati a livello aziendale, adattati alle esigenze specifiche unilateralmente o in modo bilaterale.

Nel regolamento non è attribuito all'autorità nazionale competente in materia di IA un potere approvativo dei codici di condotta, analogo a quello riconosciuto invece al Garante dalle privacy ai sensi del regolamento (UE) 2016/679. Non è escluso però che le autorità nazionali per la IA possano comunque supportare i soggetti interessati nell'adozione dei codici, sulla base del potere di orientamento e consulenza sull'attuazione del regolamento, in particolare a favore delle piccole e medie imprese, riconosciuto loro dall'articolo 70, § 8.

A ben vedere l'adozione di codici di condotta potrebbe costituire un metodo efficace di garanzia della *compliance* normativa anche nell'ambito di applicazione delle norme prescrittive del regolamento, soprattutto a favore delle piccole e medie aziende. Ciò su modello di quanto già previsto dal regolamento (UE) 2016/679 dove i codici di condotta sono concepiti come via privilegiata di conformazione normativa (articolo 40) da parte di associazioni e altri organismi rappresentanti le categorie di titolari o responsabili del trattamento per contribuire alla corretta applicazione del regolamento in specifici settori di attività e in funzione delle particolari esigenze delle micro, piccole e medie imprese. Nel contesto privacy i codici di condotta quindi assurgono a strumenti di responsabilizzazione (*accountability*), perché facilitano il rispetto delle norme e la prova della conformità. A questo riguardo si può menzionare, in campo di interesse giuslavoristico, l'esempio del codice di condotta per il settore delle APL (agenzie per il lavoro) adottato da Assolavoro e approvato dal Garante della Privacy con provvedimento 24 gennaio 2024, n. 12, che disciplina il trattamento e la conservazione digitale dei curriculum vitae dei candidati a posizioni lavorative.

Un uso analogo dello strumento dei codici di condotta può essere promosso anche nell'ambito della IA e, anzi sarebbe auspicabile che proprio attraverso i codici di condotta vengano adottate regole integrate tra regolamento (UE) 2016/679 e regolamento (UE) 2024/1689, a fine di semplificare gli adempimenti e accrescere il livello di certezza a vantaggio delle imprese e dei lavoratori ⁽³⁹⁾.

⁽³⁹⁾ Un uso dei codici di condotta è incoraggiato dal regolamento, fatta salva la natura obbligatoria e la piena applicabilità degli obblighi di trasparenza, per facilitare l'efficace attuazione degli obblighi in materia di rilevazione ed etichettatura dei contenuti generati o manipolati artificialmente, anche per sostenere modalità pratiche per favorire

Ai codici di condotta si affiancano altri strumenti di *enforcement* della regolazione europea di carattere non prescrittivo che puntano piuttosto a formare orientamenti da parte della Commissione o delle autorità indipendenti deputate alla vigilanza, a creare buone prassi da condividere e a favorire la sperimentazione normativa (c.d. *sandbox* regolativa, prevista dall'articolo 57 ss.).

A ciò si aggiungono le pratiche volontaristiche di adozione dei sistemi di gestione della qualità certificabili da enti terzi accreditati che mirano a consolidare una presunzione di conformità normativa da parte di chi se ne dota⁽⁴⁰⁾. Opera nella stessa direzione la elaborazione da parte dell'ufficio per l'IA di clausole contrattuali di tipo volontario da utilizzare tra i fornitori dei sistemi di IA ad alto rischio e altri soggetti che partecipano alla catena del valore (articolo 25, § 4).

In conclusione si può dire che sono complessivamente incoraggiate tecniche di *soft law* di varia natura, volte a promuovere un vero e proprio modello attuativo extra-legislativo del regolamento, affidato a misure incentivanti e pratiche volontaristiche al fine di innescare processi virtuosi da parte delle organizzazioni e delle imprese, secondo un approccio non vincolante, organizzativo e gestionale.

La fiducia in questo metodo è ampia al punto da affidare ad essa l'attuazione progressiva dei principi etici della IA nello spazio europeo.

È chiaro che questi sistemi si fondano sulla maturità di un tessuto di soggetti e di portatori di interessi in grado di staccarsi dalla logica del mero adempimento della norma giuridica, per abbracciare un sistema di regole convenzionali e di fonte privata. Si tratta di forme non vincolanti di conformazione dei comportamenti attesi, cui il mercato attribuisce valore, che richiedono di prendere confidenza con i sistemi di gestione della qualità e dei rischi e con il ruolo che, in tale contesto, è assunto in modo crescente dalla normazione tecnica certificabile dai soggetti accreditati.

I sistemi organizzativo-gestionali si affidano sostanzialmente a due paradigmi: l'approccio organizzativo e l'orientamento alla misurazione. Ciò implica la fissazione preventiva di obiettivi e di indicatori di performance che consentano un monitoraggio oggettivo nel tempo sul corretto sviluppo dei comportamenti attesi.

la cooperazione con altri attori lungo la catena del valore, diffondere contenuti o verificarne l'autenticità e la provenienza (cfr. considerando 135).

⁽⁴⁰⁾ È già disponibile a livello internazionale lo standard ISO/IEC 42001 per la certificazione del sistema di gestione dell'IA.

La funzione di verifica è solitamente affidata a organismi tecnici (organismi di audit) che, nel campo del lavoro, potrebbero essere opportunamente collegati alle rappresentanze sindacali. Queste ultime nei loro compiti potrebbero ben includere quello di concorrere al monitoraggio e alle valutazioni di impatto dei sistemi di IA congiuntamente con la parte datoriale. Ciò implicherebbe un salto di qualità nel ruolo sindacale nell'ambito dei sistemi di IA, evolvendo dal modello classico della informazione e consultazione, verso un nuovo modello di una partecipazione attiva di stampo organizzativo. Ecco allora che, forse, proprio su questo versante potrebbero ricavarsi spazi nuovi e costruttivi di partecipazione sindacale nella attuazione virtuosa di una IA antropocentrica e socialmente sostenibile. Uno spazio che però, allo stato, non è previsto legislativamente, ma che il sindacato potrebbe conquistare negozialmente nelle forme che ritiene più opportune ⁽⁴¹⁾.

Si tratta tuttavia di un cambiamento di portata rilevante per essere calato nella realtà concreta dei processi di gestione delle risorse umane e delle relazioni sindacali di lavoro, ma che potrebbe rappresentare una forma moderna di procedimentalizzazione dei poteri datoriali nell'epoca della IA applicata alle risorse umane.

Al cospetto delle sfide poste dalla IA in realtà è proprio l'autonomia collettiva che dovrebbe attivarsi per trovare le giuste forme di specificazione e attuazione concreta nei diversi contesti di lavoro facendosi carico di trovare nuovi equilibri e nuovi bilanciamenti. Rientra del resto a pieno titolo nella responsabilità delle parti collettive sviluppare e declinare i diritti di nuova generazione o reinterpretare principi e criteri di tutela classici del diritto del lavoro, aggiornando anche le tecniche mediante le quali adottare i controlimiti necessari a bilanciare l'esercizio dei poteri e delle prerogative datoriali.

La tecnica della procedimentalizzazione dei poteri datoriali attraverso il controllo sindacale è del resto quella più classica per il diritto del lavoro. Nel caso della IA dovrebbe puntare ad attenuare l'asimmetria informativa, accrescere la trasparenza, specialmente attraverso una partecipazione all'analisi e alla valutazione di impatto dell'uso dell'IA nei contesti di lavoro.

In alcuni Paesi europei si è optato per soluzioni più specifiche, come l'istituzione di una rappresentanza tecnica e specializzata o

⁽⁴¹⁾ L. IMBERTI, *Intelligenza artificiale e sindacato. Chi controlla i controllori artificiali?*, in *Federalismi.it*, 2023, n. 29.

l'affiancamento delle rappresentanze sindacali con figure esperte dedicate (magari con oneri condivisi o posti a carico del datore di lavoro) ⁽⁴²⁾. Senza spingersi sul terreno impervio della negoziazione o codecisione dell'algorithmo che riecheggia una concezione forte di partecipazione alla gestione dell'impresa lontana dalla tradizione nazionale di relazioni industriali, le rappresentanze sindacali potrebbero esplorare forme innovative di partecipazione inquadrando all'interno delle garanzie procedurali tipiche dei sistemi di accountability e dei modelli di organizzazione e gestione basati sull'approccio organizzativo già diffuso in campi limitrofi, come quello della salute e sicurezza, della prevenzione dei reati a vantaggio dell'impresa, della privacy e ora anche della certificazione di parità di genere. Questi modelli organizzativi sono destinati a conformare sempre più l'agire imprenditoriale, anche sulla spinta della evoluzione delle politiche di responsabilità sociale d'impresa verso forme più evolute di pratiche di sostenibilità ESG (*environmental, social and corporate governance*) e di due diligence, proiettate in una visione di filiera. Elementi essenziali di questi modelli sono la definizione da parte dell'alta direzione di politiche, strategie, obiettivi. Ma centrale, nell'ottica del miglioramento continuo, è il momento del controllo e della verifica sull'attuazione di quanto pianificato, solitamente affidati a soggetti e organismi di audit interno.

Ebbene potrebbe immaginarsi, come sviluppo futuro delle forme di partecipazione sindacale nella cornice dei modelli organizzativi fondati sulla responsabilità, un ruolo più attivo delle rappresentanze sindacali, previa formazione delle stesse, in tale contesto. Specialmente nella fase del controllo/verifica/audit nell'ambito della quale si rendono possibili valutazioni di impatto congiunte e, con esse, si creano le premesse per un contributo propositivo volto al riesame delle politiche, delle strategie e delle azioni gestionali. Con questa nuova modalità di coinvolgimento organizzativo delle rappresentanze sindacali potrebbero essere sperimentate pratiche di partecipazione sindacale orientate ad un utilizzo responsabile e sorvegliato dell'IA nei contesti lavorativi.

⁽⁴²⁾ M. CORTI, *Innovazione tecnologica e partecipazione dei lavoratori: un confronto fra Italia e Germania*, in *Federalismi.it*, 2022, n. 17; M. CORTI, *L'intelligenza artificiale nel decreto trasparenza e nella legge tedesca sull'ordinamento aziendale*, *ivi*, 2023, n. 29.

8. La via italiana alla regolazione dell'IA nel campo delle risorse umane *de iure condito* e *de iure condendo*: il puzzle normativo e il rischio della *over-regulation*

Il quadro regolatorio dell'IA si sviluppa su più livelli: concorrono a definirne i contorni fonti internazionali, dell'Unione europea e fonti interne.

A livello nazionale si è registrato un certo attivismo del legislatore. Alcuni aspetti dell'impiego dell'IA nel rapporto di lavoro sono stati disciplinati, anticipando atti regolativi europei non ancora approvati definitivamente, come il regolamento europeo sulla IA e la direttiva sul lavoro tramite piattaforma, oppure sfruttando l'occasione dell'attuazione di direttive europee a scopo più ampio, come la direttiva sulle condizioni di lavoro trasparenti e prevedibili⁽⁴³⁾. Ne viene fuori a livello nazionale un quadro composito, alquanto frammentato e certamente non organico, generato con una certa precipitosità dal legislatore anche giocando d'anticipo rispetto a provvedimenti in fase di elaborazione a livello europeo, che rischia di creare duplicazioni di norme di difficile coordinamento o, al contrario, lasciare spazi non regolati.

I principali nuclei normativi da tenere in considerazione a livello interno sono rappresentati: a) dalla disciplina in materia di protezione dei dati personali nel caso di processi decisionali automatizzati (articoli 22, 13, 14 e 15 GDPR); b) dagli obblighi informativi nel caso di utilizzo di sistemi decisionali o di monitoraggio automatizzati (articolo 1-bis, decreto legislativo n. 152/1997, modificato dal decreto legislativo n. 104/2022

⁽⁴³⁾ Per approfondimenti più specifici si veda, tra gli altri, A. TURSI, "Trasparenza" e "diritti minimi" dei lavoratori nel decreto trasparenza, in *q. Rivista*, 2023, n. 1, p. 22; M. DELFINO, *Lavoro mediante piattaforme digitali, dialogo sociale europeo e partecipazione sindacale*, in *Federalismi.it*, 2023, n. 25; D. GAROFALO, M. TIRABOSCHI, V. FILÌ, A. TROJSI (a cura di), *Trasparenza e attività di cura nei contratti di lavoro. Commentario ai decreti legislativi n. 104 e n. 105 del 2022*, ADAPT University Press, 2023; M. MARAZZA, F. D'AVERSA, *Dialoghi sulla fattispecie dei "sistemi decisionali o di monitoraggio automatizzati" nel rapporto di lavoro (a partire dal decreto trasparenza)*, in *giustiziacivile.com*, 8 novembre 2022; M. FAIOLI, *Data analytics, robot intelligenti e regolazione del lavoro*, cit.; M. FAIOLI, *Giustizia contrattuale, tecnologia avanzata e reticenza informativa del datore di lavoro. Sull'imbarazzante "truismo" del decreto trasparenza*, in *q. Rivista*, 2023, n. 1, p. 57 ss.; M.T. CARINCI, S. GIUDICI, P. PERRI, *Obblighi di informazione e sistemi decisionali e di monitoraggio automatizzati (art. 1-bis "Decreto Trasparenza")*: quali forme di controllo per i poteri datoriali algoritmici?, in *Labor*, 2023, n. 1.

c.d. decreto trasparenza); c) dagli obblighi del committente verso i lavoratori coordinati e continuativi delle piattaforme digitali (rider) previsti dall'articolo 47-*bis* ss. del decreto legislativo n. 81/2015; infine d) dalla disciplina dei controlli tecnologici verso i lavoratori subordinati previsti dall'articolo 4 Stat. lav.

L'insieme di queste norme deve ora coordinarsi con le fonti europee e, in modo particolare, con quelle del regolamento IA.

Alle difficoltà proprie della gestione di fonti multilivello, si aggiunge quella connessa al fatto che siamo in presenza di discipline a geometria variabile dal punto di vista sia soggettivo che oggettivo.

Dal primo punto di vista i soggetti protetti dalle norme sono diversi: il regolamento IA non guarda ai prestatori di lavoro, se non in materia molto limitata, guarda piuttosto a porre obblighi in capo al produttore del sistema di IA verso gli utilizzatori (*deployers*) e, a cascata, si occupa di individuare gli obblighi ricadenti in capo ai diversi soggetti che intervengono nella catena del valore dei sistemi di IA. Come visto, la prospettiva è quella della sicurezza del prodotto e del mercato unico digitale europeo. Le disposizioni del GDPR, come noto, hanno per destinatari i cittadini in generale e apprestano una tutela per gli interessati a fronte di qualsiasi trattamento di dati, indipendentemente dal fatto che si tratti di lavoratori e a prescindere dalla tipologia di rapporto che intercorre con il titolare del trattamento. Le norme del GDPR devono integrarsi, in quanto applicabili, con quelle contenute nel regolamento IA.

Sul piano interno, invece, le norme del decreto trasparenza trovano applicazione ai lavoratori subordinati e alle collaborazioni di cui all'articolo 409, n. 3, c.p.c. e a quelle etero-organizzate di cui all'articolo 2, comma 1, del decreto legislativo n. 81/2015.

Le norme sul lavoro dei rider tramite piattaforma hanno un campo di applicazione limitato ai lavoratori autonomi che svolgono attività di consegna di beni per conto altrui, in ambito urbano e con l'ausilio di velocipedi o veicoli a motore attraverso piattaforme anche digitali.

L'articolo 4 Stat. lav. si applica, come noto, soltanto al lavoro subordinato.

Anche i fenomeni regolati dalle diverse fonti menzionate sono in realtà differenti e non pienamente sovrapponibili.

Il regolamento IA si applica ai sistemi di IA come definito dall'articolo 3. La nozione legislativa italiana «di sistemi decisionali o di monitoraggio automatizzati» di cui all'articolo 1-*bis* del decreto legislativo n. 152/1997 non è precisamente definita dal legislatore interno, ma appare

sicuramente più ampia di quella dei sistemi di IA come definiti dal regolamento. Non tutti i sistemi automatizzati si configurano come sistemi di IA ai sensi del regolamento, ma soltanto quelli dotati di capacità inferenziali e gradi di autonomia nel senso sopra descritto (si veda *retro*, § 1). Viceversa non c'è dubbio che i sistemi di IA configurino *anche* sistemi automatizzati. Il che significa che ai sistemi automatizzati che non si qualificano come IA, trova applicazione la normativa italiana di cui all'articolo 1-*bis* ma non quella europea sull'IA.

La regolamentazione europea della IA trova applicazione invece a quei sistemi automatizzati che abbiano anche le caratteristiche dei sistemi di IA, ai quali dunque si applicherà in modo concorrente il regolamento IA e la normativa interna.

A sua volta la regolamentazione del GDPR troverà applicazione ai sistemi automatizzati e ai sistemi di IA soltanto se ne ricorrano i presupposti di applicazione, cioè nella misura in cui detti sistemi trattino dati personali.

Lo stesso concetto di sistema automatizzato, inoltre, non è univoco e varia nei diversi contesti regolativi, come è semplice comprendere mettendo a raffronto le diverse definizioni dei «sistemi di profilazione e decisionali automatizzati» di cui all'articolo 22 GDPR; dei «sistemi decisionali o di monitoraggio integralmente automatizzati» dell'articolo 1-*bis* del decreto legislativo n. 152/1997; dei «programmi e le procedure informatiche» della normativa sul lavoro tramite piattaforma dei rider (articolo 47-*bis*, comma 2, decreto legislativo n. 81/2015) e degli «strumenti di controllo» dell'articolo 4 Stat. lav.

Guardando all'insieme delle norme nazionali appena richiamate, pur nella diversità dei contenuti regolativi è possibile individuare alcune disposizioni ricorrenti sulle quali il legislatore ha posto l'attenzione, in una logica protettiva del lavoratore a fronte dell'impiego di sistemi per mezzo dei quali sono esercitati i poteri datoriali decisionali, organizzativi, gestionali e di monitoraggio.

Queste disposizioni ricorrenti fanno certamente perno, oltre che sull'apparato normativo della protezione dei dati personali, principalmente sull'obbligo di informazione da parte del datore/committente/titolare che radica legislativamente uno speciale diritto del lavoratore alla trasparenza algoritmica.

Nel contesto italiano l'accezione di trasparenza appare molto diversa da quella che prevale nel regolamento (UE) 2024/1689 (su cui si veda *retro*, § 6). Nell'articolo 1-*bis* del decreto legislativo n. 152/1997 l'attenzione

del legislatore interno si focalizza, facendone destinatari contemporaneamente le rappresentanze sindacali ed i singoli, tanto sulla trasparenza di interazione, quanto su una minuziosa e reiterata trasparenza di funzionamento avente ad oggetto una pluralità di dati e informazioni.

Al netto della diversa ispirazione giuridica del regolamento (UE) 2024/1689 (volto a regolare il mercato unico dei prodotti) rispetto alla normativa nazionale di stampo giuslavoristico, si può dire che l'attenzione per la trasparenza algoritmica riservata dal legislatore nazionale appare ben più ampia e pervasiva di quella dovuta in base al regolamento europeo quanto a destinatari e contenuti. Ciò non significa, tuttavia, che la via italiana alla trasparenza sia necessariamente più efficace. In particolare la trasparenza appare molto sbilanciata sulla fase iniziale e su *come fa* l'IA a pervenire alle decisioni, mentre forse sarebbe più efficace privilegiare la trasparenza di impatto, ossia una verifica *ex post* degli esiti, come punto di attenzione del processo gestionale organizzativo e delle relazioni sindacali.

Profilo, questo, che invece è quasi del tutto assente nell'articolo 1-*bis* che appare molto più interessato a far conoscere preventivamente il "come" e non il "cosa" delle decisioni assunte per mezzo della IA⁽⁴⁴⁾. Il rischio è quello di una degenerazione degli obblighi informativi che, nella moltiplicazione di informazioni tecniche scarsamente intelleggibili, rischiano di trasformare la trasparenza in una finzione giuridica, degradandola a vera e propria illusione ottica.

Se, infatti, si vuole riflettere sulla finalità ultima della trasparenza, si può concordare sul fatto che questa consiste principalmente nel mettere il singolo e suoi rappresentanti nelle condizioni di verificare un uso lecito della IA, al fine di evitare l'uso di dati non rilevanti o acquisiti in modo non legittimo; evitare l'abuso dei poteri datoriali ed evitare discriminazioni o penalizzazioni di singoli.

In ultima analisi la *disclosure* deve essere funzionale a permettere la prevenzione e la repressione degli abusi e delle distorsioni dei poteri del datore di lavoro o del diverso soggetto che applica i sistemi di IA nel mercato del lavoro.

Se questo è vero, come sembra, allora la trasparenza deve essere prevalentemente concentrata sui fattori di rischio della discriminazione e sulla pesatura dei fattori di rischio nei dati di input per garantire la qualità dei

⁽⁴⁴⁾ Su questi profili si rinvia a S. CIUCCIOVINO, *La disciplina nazionale sulla utilizzazione della intelligenza artificiale nel rapporto di lavoro*, cit.

dati in ingresso e, poi, principalmente sui risultati, soprattutto grazie al monitoraggio e alla verifica di impatto a consuntivo o *ex post*. È su questi profili che l'attenzione dei singoli e, soprattutto, delle loro rappresentanze dovrebbe focalizzarsi.

Nonostante l'alto tasso di normazione della materia della IA sia a livello europeo che a livello nazionale, il legislatore italiano è attualmente impegnato nella emanazione di una nuova legge sulla IA. È all'esame del Parlamento (XIX Legislatura) il disegno di legge AS 1146, *Disposizioni e delega al Governo in materia di intelligenza artificiale*. Secondo le intenzioni del Governo, il disegno di legge non si sovrappone al regolamento europeo sulla IA ma si propone di integrarlo per gli spazi lasciati al legislatore nazionale.

Tuttavia non si può fare a meno di evidenziare qualche criticità. A cominciare dalla definizione di IA, per la quale sarebbe stato opportuno rinviare al regolamento europeo, e dall'elencazione dei principi e delle regole che devono governare la IA, che presenta qualche imprecisione, prontamente evidenziata dal Garante della privacy nel parere reso alla Presidenza del Consiglio sulla bozza di disegno di legge. Per quanto riguarda specificamente l'applicazione della IA nel campo del lavoro (articolo 11 intitolato *Disposizioni sull'utilizzazione della intelligenza artificiale in materia di lavoro*) la formulazione della norma è tale da far sorgere il dubbio che venga delimitato drasticamente l'ambito delle applicazioni legittime della IA, circoscrivendolo alle applicazioni finalizzate a «migliorare le condizioni di lavoro, tutelare l'integrità psico-fisica dei lavoratori, accrescere la qualità delle prestazioni lavorative e la produttività delle persone» (comma 1). A di là dell'opportunità di vincolare finalisticamente i sistemi di IA a scopi stabiliti dal legislatore, in contrasto con l'ampia accezione accolta dal regolamento europeo, deve evidenziarsi che viene del tutto dimenticato il versante del mercato del lavoro e delle politiche attive dove in verità si ha un ampio utilizzo di sistemi di IA. Al punto che il richiamo del rispetto dei diritti inviolabili dell'uomo è collegato agli usi della IA nella organizzazione e gestione del rapporto di lavoro (comma 2), senza alcun accenno al mercato del lavoro.

Altre criticità sorgono con riferimento anche al richiamo all'articolo 1-bis del decreto legislativo n. 152/1997, che in parte appare superfluo (in quanto certamente applicabile ai sistemi di IA a prescindere da un suo espresso richiamo nel contesto del provvedimento) e in parte formulato in modo parziale, in quanto richiama soltanto gli obblighi informativi a

livello individuale e non quelli nei confronti delle organizzazioni sindacali (comma 3).

Da ultimo una dimenticanza, che può essere frutto soltanto di una svista del legislatore, si registra nell'ultimo comma dove si sancisce, anche qui in modo forse pleonastico, il divieto di discriminazione algoritmica, ma si tralascia di menzionare il motivo sindacale tra i fattori discriminatori (comma 3).

In considerazione della complessità e abbondanza delle norme già presenti nell'ordinamento nazionale e della diretta applicabilità del regolamento europeo, sarebbe auspicabile semmai un intervento legislativo che andasse a razionalizzare la normativa esistente e, magari, desse attuazione a quella previsione del regolamento che attribuisce agli ordinamenti nazionali la responsabilità di disciplinare i diritti di informazione e consultazione sindacale e individuale, che occorre rispettare prima di mettere in servizio o utilizzare un sistema di IA ad alto rischio sul luogo di lavoro. Tali informazioni infatti «sono fornite, se del caso, conformemente alle norme e alle procedure stabilite dal diritto e dalle prassi dell'Unione e nazionali in materia di informazione dei lavoratori e dei loro rappresentanti» (articolo 26, § 7).

In questa fase delicata di approccio ad una materia oggettivamente nuova, oltretutto attraversata da fonti normative multilivello, appare prioritario evitare il sovraccarico normativo e piuttosto optare per la definizione chiara di norme di principio e di sostegno, semmai, alla procedimentalizzazione per via sindacale dei poteri datoriali.

Il rischio è quello di una deriva iper-regolativa dell'IA che non necessariamente si traduce in un tasso più elevato di tutele per i lavoratori e le lavoratrici e che certamente non facilita il compito di chi è chiamato a mettere in atto le nuove norme al cospetto di una realtà indubbiamente molto complicata da governare.

Abstract

Risorse umane e intelligenza artificiale alla luce del regolamento (UE) 2024/1689, tra norme legali, etica e codici di condotta

***Obiettivi:** obiettivo dello scritto è quello di esaminare l'impatto del regolamento europeo sull'IA sul quadro regolativo interno nel campo delle risorse umane, tenendo conto delle applicazioni concrete dell'IA e della necessità di un'integrazione delle fonti multilivello, europee e nazionali in materia di IA. **Metodologia:** la metodologia utilizzata*

si basa su un'analisi giuridico-normativa del quadro legale europeo e nazionale con particolare riguardo alle norme giuslavoristiche. **Risultati:** la ricerca mette in evidenza i limiti derivanti dalla particolare prospettiva del regolamento (UE) 2024/1689 incentrata sull'IA come prodotto e il diverso approccio regolativo della legislazione italiana che, tuttavia, appare ancora frammentata e poco coordinata. Viene inoltre messo in evidenza il ruolo strategico dei codici di condotta, delle soft laws e delle fonti non prescrittive per favorire la più ampia diffusione dei principi etici e giuridici in materia di IA. **Limiti e implicazioni:** lo scritto mette in evidenza le implicazioni applicative del regolamento europeo e l'esigenza di riadattamento interpretativo delle fonti interne, nonché i percorsi più opportuni di implementazione della regolazione da parte dei datori di lavoro, dei deployers e delle parti sociali. **Originalità:** la ricerca fornisce uno dei primi contributi interpretativi del regolamento (UE) 2024/1689 dal punto di vista giuslavoristico e presenta una ricognizione dei casi d'uso di impiego dell'IA per le risorse umane.

Parole chiave: lavoro, intelligenza artificiale, potere digitale, trasparenza.

Human Resources and Artificial Intelligence in Light of Regulation (EU) 2024/1689, between Legal Standards, Ethics and Codes of Conduct

Objectives: the objective of the paper is to examine the impact of the European AI Regulation on the domestic regulatory framework in the field of human resources, taking into account the concrete applications of AI and the need for integration of multilevel, European and national sources on AI. **Methodology:** the methodology used is based on a legal-normative analysis of the European and national legal framework with particular regard to labour law standards. **Results:** the research highlights the limitations arising from the particular perspective of Regulation (EU) 2024/1689 focused on AI as a product and the different regulatory approach of Italian legislation, which, however, still appears fragmented and poorly coordinated. The strategic role of codes of conduct, soft laws and non-prescriptive sources in fostering the widest dissemination of ethical and legal principles on AI is also highlighted. **Limits and implications:** the paper highlights the application implications of the European regulation and the need for interpretive re-adjustment of domestic sources, as well as the most appropriate paths for employers, deployers and social partners to implement the regulation. **Originality:** the research provides one of the first interpretative contributions of Regulation (EU) 2024/1689 from a labour law perspective and presents a survey of use cases of AI use for human resources.

Keywords: labour, artificial intelligence, digital power, transparency..