

Smart Factory, dignità del lavoratore ed intelligenza artificiale come forma di autoapprendimento¹

Valerio Maio

1. Civiltà e inciviltà digitale

Nel ringraziare dell'onore immeritato di questo invito, non posso non indugiare brevemente sul filo conduttore opportunamente suggeritoci dal Prof. Giuseppe Santoro Passarelli per legare fra loro le riflessioni di quest'oggi, ovvero il rapporto tra «civiltà digitale» e «dignità» di chi lavora.

L'espressione «civiltà digitale» è particolarmente felice perché fotografa come le nuove tecnologie abbiano ormai colonizzato ogni ambito, andando a costituire un vero *habitat* sociale, dove non solo il lavoro può essere *smart*, ma la casa diviene *smart*, l'auto deve essere *smart*, la città va ripensata come una *smart city*. Se assecondiamo il paradigma della cd. «infosfera»², tutti gli ambienti di vita e lavoro devono essere congegnati per consentire alle nuove tecnologie digitali una diffusione capillare e favorirne l'incondizionata espressione.

D'altra parte, il binomio «civiltà digitale» rievoca alla mente la formula «civiltà industriale» con cui additiamo un altro modello di organizzazione della

¹ Il presente scritto costituisce la rielaborazione della relazione svolta nell'ambito del Convegno *Dignità del lavoro e Civiltà digitale*, organizzato dall'Accademia Nazionale dei Lincei, tenutosi in Roma il 24 febbraio 2023, e sarà pertanto pubblicato anche nel volume che ne raccoglie gli Atti.

² Che secondo Floridi 2022 sintetizza il fenomeno del cd. «avvolgimento» del mondo.

Valerio Maio, Sapienza University of Rome, Italy, valerio.maio@unitelmasapienza.it, 0000-0002-2144-2268

Referee List (DOI 10.36253/fup_referee_list)

FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Valerio Maio, *Smart Factory, dignità del lavoratore ed intelligenza artificiale come forma di autoapprendimento*, © Author(s), CC BY 4.0, DOI 10.36253/979-12-215-0507-8.48, in William Chiaromonte, Maria Luisa Vallauri (edited by), *Trasformazioni, valori e regole del lavoro. Scritti per Riccardo Del Punta*, pp. 783-803, 2024, published by Firenze University Press, ISBN 979-12-215-0507-8, DOI 10.36253/979-12-215-0507-8

realtà, dove il processo di produzione della ricchezza storicamente veniva delegato alle fabbriche, capaci di catalizzare la forza lavoro attiva dell'intera società.

Solo che il lemma «civiltà», presupposto in entrambi i casi, non ci deve ingannare, perché non è affatto scontato che la diffusione pervasiva di una tecnologia produca anche una civilizzazione della realtà³.

Quella della civilizzazione del mondo digitale è, a ben vedere, una partita ancora tutta da giocare, il cui esito dipende in larga parte dalla competenza e dal coraggio dei decisori politici, ma anche dalla consapevolezza dell'opinione pubblica e, dunque, al fine, per quanto in *parva pars*, anche dalla qualità del nostro dibattito odierno.

Dibattito nel quale, mi pare di poter dire, percepiamo tutti il rischio dell'affermarsi di una inciviltà digitale. Come era incivile la condizione in cui versavano all'alba della prima industrializzazione i lavoratori che Dickens, nei *Tempi difficili*, chiamava semplicemente «le mani».

Anche ragionando dei nostri tempi difficili dovremo avere presente la sottile analisi dickensiana, che denunciava come, nel mentre tutta l'organizzazione sociale si modellava acriticamente a misura delle esigenze della fabbrica, il nuovo *habitat* di lavoro produceva conseguenze irreversibili sulla dignità di chi vi lavorava.

Certamente possiamo convenire che il livello della protezione giuridica del lavoro attuale non è neppure paragonabile a quella di cui beneficiavano i lavoratori della prima o seconda industrializzazione. Ma lo stesso divario non può non essere misurato anche con riferimento al livello di sofisticatezza, diffusione ed accettazione sociale che contraddistingue le nuove tecnologie, se anche solo per un istante le poniamo impietosamente a confronto con un telaio a vapore.

Onde, non mi pare dubitabile che, così come qualcosa di realmente mai visto si trasmetteva direttamente dalle «mani» alle «vite» dei nostri antenati, anche l'introduzione massiccia ed incontrollata nella fabbrica cd. intelligente di dispositivi tecnologici potenziati da intelligenza artificiale potrebbe avere conseguenze dirette sulla dignità di chi vi lavora.

2. Dispositivi capaci di auto-apprendimento non supervisionato nei luoghi di lavoro

Che il lavoro in ogni ambito conviva ormai da decenni con l'impiego sempre crescente di innovazioni tecnologiche digitali è, peraltro, fatto notorio. In taluni settori merceologici, come ad esempio la logistica⁴, grazie alla intelligenza

³ Sull'utilizzo spesso ambiguo del termine, Mazlish 2004.

⁴ In realtà, ovunque il *problem solving* è affidato all'intelligenza artificiale si hanno ricadute sul lavoro. Un esempio concreto può essere la gestione dei porti, dove oggi l'intelligenza artificiale guida la movimentazione delle navi in passato affidata all'uomo, impiegando sensori intelligenti ed algoritmi di analisi dati che, non solo assicurano una maggiore sicurezza ed una migliore efficienza nella movimentazione delle merci, ma sanno anche predire eventuali congestionamenti del traffico, così da ovviare a ritardi che possono tradursi in perdita economica.

artificiale si tende dichiaratamente alla marginalizzazione del lavoro umano, procedendo spediti verso la fabbrica «a luci spente».

Certamente anche simili processi di marginalizzazione dell'apporto umano sono in grado di produrre conseguenze sulla dignità del lavoro⁵, perché nella logica di mercato la debolezza del lavoro dipende anche dal reciproco dimensionamento di domanda ed offerta di lavoro. Ma si tratta di conseguenze indirette che, per questo, dobbiamo oggi accantonare al di fuori dell'ambito di questa relazione.

Il nostro tema è infatti la fabbrica, ma anche l'ufficio, nella quale operano, fianco a fianco con i lavoratori, sistemi di intelligenza artificiale cd. forte, divincolata cioè dalla meccanica deterministica *input/output* ed in grado piuttosto di auto-apprendere ed auto-determinarsi.

L'ingresso nei luoghi di lavoro di questi sistemi avviene sia per il tramite della robotica potenziata dall'intelligenza artificiale, sia attraverso l'impiego di programmi o *software*, come *bot* o *chatbot*, privi di un riconoscibile supporto fisico. Perché, lo ha compreso oramai anche il regolatore euro-unitario, nella *smart factory* digitale non conta la presenza oggettuale, ma il modello di intelligenza algoritmica che viene adottato⁶.

Più ancora in particolare, il titolo che mi è stato consegnato, se bene l'ho inteso, pone l'accento sull'intelligenza artificiale capace di apprendimento non supervisionato, che opera dunque autonomamente, confrontando i dati e ricercando similarità o differenze nelle sequenze di dati (cd. *pattern*).

Il *machine learning* non supervisionato è infatti la tecnica di apprendimento automatico che consiste nel fornire una serie di *input* che il sistema informatico riclassificherà ed organizzerà autonomamente sulla base delle caratteristiche comuni rilevate, per cercare di effettuare ragionamenti e previsioni quanto agli *input* successivi (di qui quella che chiamiamo la capacità predittiva) (Cristianini 2023).

In sostanza, quando assumiamo che vi sia una non meglio precisata autonomia del sistema intelligente che interagisce con i lavoratori, il più delle volte stiamo al fondo dicendo che degli algoritmi di apprendimento non supervisionato cercano di scoprire *pattern* nascosti nei dati, con l'obiettivo di raggruppare, classificare, separare o manipolare quei dati, e così, ad es., predire quale sia un output accettabile od identificare una anomalia rispetto ad uno standard. L'autonomia risiede dunque in ciò che l'apprendimento non viene supervisionato⁷

⁵ Colaiacomo 2022, 211, inquadra la perdita del lavoro per sostituzione tecnologica come «perdita della base della dignità del lavoratore».

⁶ Era invece questo il fraintendimento ed il limite più evidente del primo tentativo europeo di disciplinare la robotica potenziata da intelligenza artificiale effettuato con la Risoluzione del Parlamento dell'Unione europea del 16 febbraio 2017, su cui in senso critico Maio 2018a, sp. 1424 sgg.

⁷ Diversamente nell'apprendimento automatico cd. supervisionato l'intelligenza artificiale viene allenata sulla base di una serie di esempi, costituiti da coppie ideali di *input* e di *output*. Un tipico esempio di *set* dati di allenamento è l'insieme delle cartelle cliniche oggetto di diagnosi umana nel passato, che viene processato per istruire l'intelligenza artificiale a

ed i dati di *input* non sono stati prima etichettati o non hanno un corrispondente valore di *output* predefinito, per cui il sistema apprende direttamente dalla propria «esperienza», o meglio da una sua traduzione in dati, senza avere esempi o risposte di riferimento⁸.

Questo non significa affatto che la «mano» dell'uomo scompaia del tutto. Per potere raggiungere l'obiettivo prefissato, l'algoritmo che apprende autonomamente deve, infatti, prima essere addestrato, mediante un *training* basato su di una serie di esempi, a loro volta non etichettati, rispetto ai quali impara ad individuare i *pattern* che legano i dati, ad esempio muovendo da ipotesi preconfigurate e deducendo le probabilità che un evento accada alla ricorrenza di determinate variabili.

L'addestramento dell'algoritmo viene quindi controllato nella fase di *training* attraverso gli *iperparametri*, che permettono di tarare il modello che interpreta i dati. In relazione ai valori di *output* è così possibile disegnare, sia pure a grandi linee, la logica di funzionamento del sistema intelligente considerato.

Di fronte all'ingresso sui luoghi di lavoro di questi modelli di intelligenza artificiale diviene indispensabile focalizzare quali problemi inediti sorgono per il diritto del lavoro.

E mi pare che, allo stato delle nostre conoscenze, ne possiamo già evidenziare almeno tre.

Abbiamo anzitutto il problema delle cd. decisioni datoriali algoritmiche non spiegabili. Dove l'inspiegabilità, che può sembrare un paradosso per la lettura classica dell'impresa come soggetto razionale, viene alimentata dall'imperscrutabilità e dall'opacità degli algoritmi medesimi.

Abbiamo poi il tema dei cd. *bias* algoritmici che inducono discriminazioni al momento della selezione e poi gestione del personale, peraltro in molti casi è stato rilevato proprio a causa del rinforzo umano nella fase di apprendimento.

Abbiamo infine l'emergere di una capacità predittiva e decisionale di fatto imponderabile, che agisce all'interno di una relazione di potere fra esseri umani costitutivamente asimmetrica come è quella di lavoro, che fino ad oggi gestiva-

riconoscere una determinata patologia. L'apprendimento non supervisionato, quello tipico ad esempio dei motori di ricerca, a sua volta, può essere di tipo cd. rinforzato, quando si avvale di modelli ottimizzati tramite il rinforzo umano inconsapevole, secondo la tecnica nota come *Reinforcement Learning From Human Feedback*. Quando cioè l'intelligenza artificiale continua ad accrescere la propria competenza e ad affinare la propria sensibilità apprendendo autonomamente dai *feedback* degli utenti.

⁸ Al contrario degli algoritmi espliciti che dipendono deterministicamente dal dato immesso, i sistemi informatici che implementano algoritmi di *machine learning* possono essere descritti come autonomi o semiautonomi nella misura in cui i loro risultati sono solo indotti dai dati in senso non deterministico. Questi sistemi infatti producono *output* su base di una valutazione probabilistica, che si fonda esclusivamente su ragionamenti e previsioni statistiche rispetto agli *input* successivi. Per una disamina di queste distinzioni v. già la Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi, adottata dalla Commissione Europea per l'efficienza dei sistemi giudiziari e negli ambiti connessi, XXXI Riunione plenaria, Strasburgo, 3-4 dicembre 2018 e cfr. Peruzzi 2023, sp. 11 e sgg.

mo per mezzo di valutatori umani, che si riferivano a prescrizioni e clausole generali destinate ad essere interpretate dall'intelligenza umana, in applicazione di una logica valoriale umana⁹.

Nel complesso, il diritto si trova dunque urgentemente a fare i conti con una strisciante preordinata disumanizzazione del lavoro (Gragnoli 2020).

Proviamo a distinguere in quali ambiti e con quali ricadute.

3. *Cobot* intelligenti e potenziamento del lavoro umano

L'intelligenza artificiale è presente nei luoghi di lavoro in primo luogo per la crescente diffusione dei *cobot* o robot collaborativi, programmati non più solo per svolgere meccanicamente un determinato compito, ma per interagire con l'uomo, apprendendo nella e dalla interazione con i lavoratori, sempre più spesso senza barriere protettive rigide, all'interno delle cd. zone di sicurezza dinamiche.

I *cobot* resi autonomi dall'intelligenza artificiale sono stati uno dei pilastri di una recente stagione di aggiornamento del nostro sistema produttivo, passata alle cronache con la formula «industria 4.0» (Del Punta 2018). Rimangono ad esempio centrali in tutte le attività di assemblaggio collaborativo, dove vengono impiegati per elevare le prestazioni di processo.

In realtà, il *cobot* prelude ad un cambio di paradigma nella stessa idea di fabbrica. Se il robot aziendale classico era una macchina deputata a svolgere con velocità e precisione inumane compiti massivi e ripetitivi, quasi sempre a costi proibitivi, che richiedevano perciò investimenti pluridecennali; il *cobot* invece è progettato per apprendere dall'uomo, con cui collabora anche grazie a strumenti di visione artificiale e sensoristica avanzata, una volta attivata la modalità *learn* o *training*, e deve, a costi accessibili anche alle piccole e medie imprese, garantire flessibilità ed adattamento ai molteplici processi nei quali può essere di volta in volta inserito.

Con riferimento ai *cobot* intelligenti, il principale problema è relativo alla sicurezza della relazione tra l'uomo e la macchina intelligente.

Al riguardo, già in altra occasione ho osservato che, rispetto ad una fabbrica tradizionale in una *smart factory*, di norma, il lavoro dovrebbe essere più sicuro. Perché statisticamente, sia pure con differenze da settore a settore, l'esperienza è che l'incidenza degli infortuni decresce con l'ammodernamento tecnologico dell'impresa¹⁰.

⁹ Per la constatazione solo apparentemente scontata che «il diritto è umano, fatto dagli esseri umani e per gli esseri umani», Pietropaolo 2022, 39.

¹⁰ Maio 2018a, 1427 sgg. Allo stesso tempo, osservammo che nel diritto del lavoro non si registra un dibattito paragonabile a quello che sta impegnando a fondo il diritto civile in merito ai profili della responsabilità per l'impiego di intelligenza artificiale (cfr. Costanza 2018 e da ultimo Peruzzi 2023, 131 sgg.). Questo si deve al fatto che nel diritto del lavoro opera l'art. 2087 c.c. e, dunque, anche nella *smart factory* il datore di lavoro risponde dell'infortunio cagionato dal malfunzionamento di un *cobot*, senza necessità di deviare dai consueti schemi basati sul rischio ambientale o sulla presunzione di pericolosità delle macchine mosse.

Certo il *cobot* reso autonomo dall'intelligenza artificiale e deputato ad interagire con l'uomo ci allarma più e diversamente rispetto agli altri dispositivi non autonomi, perché potrebbe dare causa ad eventi a maggior ragione non preventivabili neppure per il datore di lavoro (è il tema dei cd. comportamenti emergenti).

Ma anche qui, l'effettiva pericolosità dell'interazione del *cobot* con l'uomo dipende dalla qualità del *design* informatico, che ne definisce le possibilità di azione ed il grado di reale autonomia, e dal sistema prevenzionistico in cui è inserito, oltre che dal corretto o scorretto impiego che ne fa l'uomo¹¹.

Il punto è allora che l'inserimento del *cobot* all'interno della fabbrica intelligente richiede una specifica ridefinizione, e forse anche più di qualche adattamento, dello strumentario giuridico tradizionale. Impone una specifica informazione e formazione dei lavoratori chiamati ad interagire con il *cobot*. Richiede la mappatura e valutazione dei nuovi rischi derivanti dalla interferenza tra l'uomo e i dispositivi intelligenti, come anche dalla eventuale interrelazione fra dispositivi autonomi rispetto all'uomo. Oltre che dei rischi propriamente riconducibili all'impiego di forme di intelligenza artificiale connessioniste che apprendono autonomamente, tra cui il cd. *catastrophic forgetting* (Ans, Rousset, French, Musca 2004). Implica inoltre, sempre in ottica precauzionale, l'adozione di *policy* rigorose di *cyber security* e di un *disaster recovery plan*¹².

Problemi molto complessi derivano poi dall'impiego nella fabbrica intelligente di strumentazioni potenziate dall'intelligenza artificiale applicabili direttamente sul corpo del lavoratore, come arti intelligenti o esoscheletri digitali.

La comparsa sui luoghi di lavoro di lavoratori aumentati, anche se pone interrogativi difficili (Maio 2020a, sp. 513 sgg.), promette ricadute positive.

Va dato atto, ad esempio, che nel settore dell'*automotive* già oggi gli esoscheletri intelligenti coadiuvano i lavoratori impiegati in lavori usuranti, come quelli cd. sottoscocca, per sostenerne la muscolatura e la tenuta ossea, evitando infortuni e prevenendo le malattie professionali.

Di più, il potenziamento tecnologico può garantire un'accessibilità al mercato lavoro anche ai lavoratori penalizzati da deficit fisici preesistenti che altrimenti ne rimarrebbero esclusi.

¹¹ In tema di eziologia infortunistica, si discute dell'apporto causale o concausale del formatore del *cobot* durante le sessioni di *training*. Se cioè l'immissione di dati scorretti, o comunque altre forme di sviamento del processo di apprendimento dell'intelligenza artificiale, possono far ritenere il lavoratore incaricato di formare il *cobot* corresponsabile dell'infortunio, in ragione peraltro di un inadempimento vero e proprio, prima ancora che per una qualche *culpa in educando*. Ma anche qui, dobbiamo rifuggire inutili allarmismi. Ed assumere che di norma la fabbrica intelligente è piuttosto una realtà organizzata e controllata, dove operano lavoratori fortemente specializzati e intensamente formati. Mentre, a sua volta, il *cobot*, prima ancora di venire inserito in un simile contesto, è già stato testato e tarato dal produttore su *set* di dati controllati. Se così è, lo sviamento dell'apprendimento per essere imputabile al lavoratore e soprattutto condurre alla irresponsabilità del datore di lavoro dovrebbe essere doloso, o se mai assumere i connotati della deliberata assunzione di un rischio elettivo.

¹² Su tutti questi aspetti cfr. ancora se vuoi Maio 2018a, sp. 1430 sgg.

Allo stesso tempo, però, tutte le tecnologie che producono un *enhancement* del lavoratore vanno indagate a fondo per le possibili implicazioni sulla libertà di scelta di chi lavora, affiancando a quella consueta sulle ricadute di sicurezza (ad es. del carico biomeccanico) anche una rigorosa valutazione dell'impatto psicologico e sulla personalità morale del lavoratore di questi dispositivi.

Proteggere la dignità di chi lavora significa anche consentire l'ingresso nei luoghi di lavoro soltanto a quei dispositivi intelligenti che si dimostrano utili per facilitare o potenziare l'espressività del lavoro umano. E di converso impedire all'intelligenza artificiale di prevaricare la libera determinazione della persona, ad esempio, spingendosi a condizionarne i movimenti e la comunicazione o a dettare le cadenze ed i tempi di risposta con una invasività tale da deumanizzare i comportamenti o la gestualità del lavoratore.

Potremmo forse racchiudere questo principio di cautela con una formula generale: l'ordinamento può consentire al lavoratore, sia pure entro limiti e con precauzioni da stabilire caso per caso a tutela della sua salute e personalità, di indossare dispositivi che impiegano l'intelligenza artificiale per aumentare le potenzialità dell'umano. Ma non mai viceversa.

4. Gestione del personale ed esercizio dei poteri del datore di lavoro mediante intelligenza artificiale

L'intelligenza artificiale quando si traduce in decisioni algoritmiche automatizzate ha ormai un ruolo determinante anche nell'organizzazione e gestione del lavoro.

A partire dalla selezione dei candidati, dove l'intelligenza artificiale è impiegata nella fase di costituzione del rapporto di lavoro per lo *screening* curriculare. Ma anche con riferimento allo svolgimento del rapporto di lavoro, quando la cd. pianificazione intelligente può dettare le scelte aziendali strategiche e di riorganizzazione, o persino determinare in concreto l'esercizio dei poteri datoriali, direttivo, di controllo e disciplinare.

In una sorta di rovesciamento della relazione tra intelligenze, quella artificiale sembra ormai pronta per eterodirigere in tempo reale l'uomo che lavora, controllarlo e valutarlo, e poi incentivarlo o sanzionarlo, in relazione agli obbiettivi aziendali.

Nella prospettiva della cd. gestione algoritmica dei rapporti di lavoro, i sistemi decisionali e di monitoraggio automatizzati basati su algoritmi, con la promessa di efficientarla, esautorano di fatto la direzione aziendale.

Ma proprio la prospettiva del cd. *management* algoritmico presuppone, anzitutto, che venga stabilito quando si realizza una valutazione algoritmica. Conclusione che non è affatto semplice, specie se si considera il ritmo di diffusione e la continua evoluzione di questi strumenti ed applicativi.

Una prima risposta è stata data dal nostro ordinamento con riguardo al modello lineare di gestione algoritmica del lavoro che è tipico delle piattaforme digitali, dove, se stiamo alla declinazione per i cd. *riders* (Santoro Passarelli 2022), che ha fin qui catalizzato l'attenzione anche dello stesso legislatore, l'intelligenza artificiale gestisce le chiamate, indica i percorsi, valuta le prestazioni, distribuisce premialità o sanzioni.

Ma valutazioni algoritmiche sono oggi possibili praticamente in ogni ambito e momento dell'attività lavorativa e con modalità non sempre tracciabili. E questo rende anche la questione qualificatoria più sfumata e complicata.

Per intenderci può essere utile fare un esempio: una casa editrice può oggi valutare il lavoro dei propri editor confrontando le sinossi presenti nelle quarte e bandelle dei volumi pubblicati nell'ultimo anno, con quelle realizzate in pochi secondi, con un semplice comando vocale, dal programma «ChatGPT».

In questo caso, è oggettivamente problematico stabilire se siamo ancora nell'ambito di una valutazione umana, per quanto assistita dall'intelligenza artificiale, o di fronte ad una vera e propria valutazione algoritmica.

La risposta sarebbe invece scontata se ipotizzassimo che il sistema dotato di intelligenza artificiale provvede direttamente da sé a confrontare e valutare quelle sinossi, per poi classificare gli editor sulla base di parametri introiettati nella fase di *training* ovvero auto-definiti apprendendo dai dati processati, ed infine procedere alla scelta di chi premiare, piuttosto che licenziare.

Perché in questa seconda ipotesi, anche a dispetto del medesimo risultato finale, l'intelligenza artificiale perde il connotato della strumentalità per assumere le spoglie del datore di lavoro ed impugnare direttamente le redini della relazione giuridica di lavoro.

Ora, rispetto a questi scenari, l'interprete deve anzitutto prendere atto che il legislatore comunitario e quello nazionale stanno fornendo indicazioni inequivocabili nel senso di ritenere la gestione algoritmica dei rapporti di lavoro certamente ad alto rischio, ma allo stesso tempo pienamente legittima.

Mentre in sede euromunitaria si avvia alla fase decisiva la discussione sul quadro regolatorio basato sulla valutazione del rischio e sul principio che un sistema di intelligenza artificiale deve essere vietato solo se si dimostra portatore di pericoli inaccettabili¹³, non si vede all'orizzonte alcuno che sostenga, neppure dal versante sindacale, le ragioni di un divieto generalizzato di gestione algoritmica del lavoro.

Divieto che, evidentemente, viene considerato, prima ancora che antistorico ed antieconomico (in termini di isolamento tecnologico), assai difficile da far rispettare. Visto che, al fondo, basterebbe il ricorso alla ratifica formale delle decisioni algoritmiche da parte di decisori umani per produrre un qualche aggiramento di un eventuale divieto¹⁴.

La partita regolatoria si sposta non a caso sulle condizioni di esercizio del *management* algoritmico, dove registriamo un primo intervento norma-

¹³ Cfr. la Proposta di Regolamento del parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale, Bruxelles, 21.4.2021 COM(2021) 206 final, su cui Alpa 2022. *L'Artificial Intelligence Act*, in data 14 giugno 2023 ha come noto avuto il via libera del Parlamento europeo.

¹⁴ Occorre sempre considerare che la decisione algoritmica non è in quanto tale, per definizione, necessariamente migliore o peggiore, e neppure maggiormente penalizzante o più inaccettabile, per chi lavora, rispetto a quella invece imputabile al datore di lavoro che non si affida all'intelligenza artificiale.

tivo di *hard law*, con l'art. 4, co. 1, lett. b, del decreto legislativo 27 giugno 2022, n. 104¹⁵.

5. Il decreto legislativo 27 giugno 2022, n. 104 ed il contrasto all'opacità algoritmica

Questa previsione è, in senso proprio, epocale.

Perché certifica che il tema non è stabilire se sia possibile utilizzare sistemi decisionali o di monitoraggio automatizzati che gestiscono il rapporto di lavoro, ma soltanto a quali condizioni è legittimo farlo.

Senza forse essercene accorti, abbiamo tutti così valicato un confine. E «oltre» quel confine si dischiudono possibilità imprevedibili fino a soltanto pochi anni fa.

Anche qui un esempio può aiutare a chiarire. Il legislatore riconosce così che anche un tema sensibile come la morbilità aziendale possa venire gestito per il tramite di valutazioni e determinazioni dell'intelligenza artificiale.

E del resto non v'è nulla da obiettare se il calcolo del cd. comporta per malattia viene demandato ad un *software* automatizzato, capace di stabilire se e quando vi è stato superamento del termine di sospensione, integrando la disciplina sindacale applicabile (a dispetto delle molteplici interpretazioni cui le norme dei contratti collettivi si prestano) con un puntuale ed aggiornato monitoraggio degli orientamenti giurisprudenziali. L'intelligenza artificiale si sostituirebbe al fondo al consulente aziendale.

Ma, allo stesso tempo, penso che avvertiremmo tutti l'inadeguatezza dello strumento rispetto allo scopo se l'intelligenza artificiale venisse direttamente investita anche della decisione di licenziare il lavoratore affetto da eccessiva morbilità.

Perché chiunque ha esperienza di direzione del personale sa che una decisione di questo tipo necessita di sensibilità, pensiero laterale, capacità di capire il contesto entro cui si cala la vicenda personale, prima ancora che lavorativa, del dipendente considerato, che non può essere trattato semplicemente come dato da processare e la cui sorte non è riducibile a mero *output*.

¹⁵ Il cd. decreto trasparenza, in attuazione della Direttiva UE 2019/1152 del Parlamento europeo e del Consiglio del 20 giugno 2019, relativa alle condizioni di lavoro trasparenti e prevedibili, ha novellato il d.lgs. 152 del 1997 introducendo un art. 1 bis giusta il quale «Il datore di lavoro o il committente pubblico e privato» sono obbligati «a informare il lavoratore dell'utilizzo di sistemi decisionali o di monitoraggio integralmente automatizzati» se questi sistemi sono «deputati a fornire indicazioni rilevanti ai fini della assunzione o del conferimento dell'incarico, della gestione o della cessazione del rapporto di lavoro, dell'assegnazione di compiti o mansioni», nonché se quei sistemi sono deputati a dare «indicazioni incidenti sulla sorveglianza, la valutazione, le prestazioni e l'adempimento delle obbligazioni contrattuali dei lavoratori». L'importanza della disposizione è stata presto colta dalla dottrina che l'ha commentata per lo più con accenti critici, cfr.: Marazza, D'Aversa 2022, Proia 2022a, Proia 2022b, Garofalo, Tiraboschi 2023, Schiavone 2023, Carinci, Giudici, Perri 2023, Faioli 2023, Zilli 2023.

L'art. 4, co. 1, lett. b, del decreto legislativo n. 104 cit. è una norma estremamente importante anche per un altro motivo.

Sulle orme di quanto già introdotto in ambito *privacy* con l'art. 22 del Regolamento dell'Unione europea n. 679/2016¹⁶, esplicita anche in territorio strettamente giuslavoristico un primo fondamentale presidio di garanzia che consiste nel diritto dei lavoratori di sapere, e di conseguenza nell'obbligo del datore di lavoro di informare, se una decisione che impatta sul rapporto di lavoro è ascrivibile alla valutazione dell'intelligenza artificiale.

Come noto, dopo un discusso intervento correttivo, la disposizione precisa ora il proprio campo di applicazione con riguardo all'utilizzo di sistemi decisionali o di monitoraggio «integralmente» automatizzati, e non più solo genericamente automatizzati¹⁷.

La correzione è stata viepiù interpretata come un tentativo di ridimensionare la portata applicativa della norma. Ma la questione, almeno a mio avviso, non muta poi di molto. Come abbiamo visto con l'esempio degli editor e delle sinossi, è difficile soppesare in termini quantitativi il grado di automazione di un processo decisionale partecipato dall'intelligenza artificiale.

Diversamente, tanto prima quanto dopo l'intervento correttivo, ciò che rileva ai fini della applicazione, o no, dell'obbligo di informativa è piuttosto l'elemento qualitativo della vincolatività della decisione algoritmica una volta prodotta dal sistema automatizzato.

Se, cioè, la stessa semplicemente contribuisce alla corretta e completa informazione della direzione aziendale, prodromica rispetto ad una determinazione che resta, nella sostanza e non solo formalmente, imputabile ad una volontà umana (anche se oggettivata in forma collettiva). Nel qual caso non avrebbe senso l'obbligo di informativa. O, piuttosto, sia invece capace di incidere direttamente in maniera precettiva, e non semplicemente strumentale, sulla vita lavorativa del dipendente¹⁸.

Anche ove si consolidasse questa interpretazione che riconsegna alla norma spazi ragionevoli di concreta operatività, la regolazione introdotta con il decreto

¹⁶ Sui limiti di impiego dell'art. 22 cit. nel rapporto di lavoro Gagnoli 2022, sp. 35 sgg.

¹⁷ Così dopo la modifica recata dall'art. 26, co. 2, lett. a), del d.l. 4 maggio 2023, n. 48, convertito con modificazioni dalla l. 3 luglio 2023, n. 85. La formulazione originaria sembrava potere avere un campo di applicazione più ampio anche rispetto all'art. 22 del GDPR, tale da includere qualsivoglia determinazione aziendale che avesse, sia pure a valle del processo decisionale, beneficiato dell'impiego di intelligenza artificiale. Per tentare di arginare questo tipo di interpretazioni «neoluddiste» il Ministero del lavoro è corso ai ripari con la circolare n. 19 del 20 settembre 22, che, se da un lato ha così confermato la plausibilità delle letture osteggiate, dall'altro ha finito col fare riferimento ad un criterio di «accessorietà» dell'intervento umano che la norma, invero, non contemplava.

¹⁸ Questa lettura sembrerebbe confermata anche la prima giurisprudenza in tema, secondo la quale, per escludere che un sistema algoritmico possa ritenersi integralmente automatizzato, non sarebbe sufficiente la possibilità in astratto di un intervento umano, ma occorrerebbe dimostrare la necessità dell'intervento umano ed il fatto che l'intervento umano, oltre a ricorrere nella generalità dei casi, non sia meramente formale, così Trib. Torino 7 agosto 2023 in Wikilabour.it.

legislativo n. 104 cit. potrebbe lo stesso apparire velleitaria, visto che al fondo oppone un baluardo apparentemente emaciato come il principio di trasparenza alla straordinaria invasività dei sistemi decisionali e di monitoraggio automatizzati nel rapporto di lavoro.

In realtà, a dispetto di un approccio regolatorio che certamente tradisce una ascendenza molto più algoretica che giuridica, la disposizione non va sottovalutata. Anzi, proprio perché la trasparenza costituisce la precondizione minima per ragionare di equità algoritmica, sarebbe probabilmente miope predicarne l'irrelevanza nel diritto del lavoro.

L'opacità algoritmica del resto conduce all'inspiegabilità dei processi decisionali, che è alla base dei *bias* e delle discriminazioni algoritmiche, quando non favorisce il cd. *blue washing*, ossia la possibilità di edulcorare, avvalendosi del pregiudizio positivo per l'automazione, comportamenti non etici, che, però, nel caso del rapporto di lavoro sono, il più delle volte, anche potenzialmente illeciti od illegittimi (pensiamo alla selezione curriculare automatizzata i cui esiti rispondono a pregiudizi razziali o di genere introiettati dall'intelligenza artificiale).

Dunque, l'obiettivo di progredire verso una intelligenza artificiale per quanto possibile trasparente e in qualche misura spiegabile, che accomuna l'etica e il diritto, passa anche dalla capacità di articolare in maniera puntuale e pertinente i contenuti dell'informativa al lavoratore¹⁹.

Quanto poi alla capacità di queste previsioni di dimostrarsi effettive, la valutazione dipenderà in ampia parte dal grado di dettaglio e di approfondimento tecnico che la giurisprudenza valuterà obbligatorio per le informative.

Se cioè prevarrà la logica del mero adempimento formale, del rinvio generico a formule standard preimpostate, per cui sarà sufficiente etichettare la famiglia di algoritmi impiegati come «non deterministici», «evolutivi», «capaci di apprendimento non supervisionato». Od invece, verrà imposto al datore di lavoro di spingersi fino a dettagliare la logica ed il funzionamento dei sistemi decisionali o di monitoraggio automatizzati che utilizza nella propria azienda.

Certo, per le ragioni viste prima (cfr. supra il paragrafo n. 2), quando si ha a che fare con un'intelligenza artificiale capace di autoapprendimento continuo e

¹⁹ Le informazioni considerate dall'art. 1 bis del d.lgs. n. 152 del 1997 cit. riguardano: «a) gli aspetti del rapporto di lavoro sui quali incide l'utilizzo dei sistemi decisionali o di monitoraggio automatizzati; b) gli scopi e le finalità dei sistemi decisionali o di monitoraggio automatizzati; c) la logica ed il funzionamento dei sistemi decisionali o di monitoraggio automatizzati; d) le categorie di dati e i parametri principali utilizzati per programmare o addestrare i sistemi decisionali o di monitoraggio automatizzati, inclusi i meccanismi di valutazione delle prestazioni; e) le misure di controllo adottate per le decisioni automatizzate, gli eventuali processi di correzione e il responsabile del sistema di gestione della qualità; f) il livello di accuratezza, robustezza e cybersicurezza dei sistemi decisionali o di monitoraggio automatizzati e le metriche utilizzate per misurare tali parametri, nonché gli impatti potenzialmente discriminatori delle metriche stesse».

non supervisionato, anche solo stabilire quale sia la logica di fondo od il criterio di funzionamento non è affare semplice e può risultare frustrante²⁰.

Perché la caratteristica intrinseca di queste famiglie di algoritmi è proprio quella di alterare la propria logica decisionale durante il processo di autoapprendimento, e di farlo con una frequenza ed una velocità tale che rende molto complesso, e forse ai limiti del possibile, ricostruire la catena che dalla cognizione iniziale dei dati porta ad una decisione. Circostanza che si traduce in quella incapacità di aprire la cd. *black box* algoritmica dietro cui si schermano gli informatici (Pasquale 2016).

Ma anche queste innegabili difficoltà non hanno evidentemente convinto il legislatore a rassegnarsi ad accettare remissivamente la convivenza nei luoghi di lavoro con forme tiranniche di intelligenza artificiale o volontà algoritmiche aprioristicamente incomprensibili²¹.

Dei molti ostacoli che si frappongono rispetto al traguardo di una reale trasparenza²² si mostra del resto consapevole lo stesso decreto n. 104 cit., visto che, opportunamente, applica l'obbligo di informativa non ai dati, ma alle «categorie di dati», non a tutti i parametri, ma soltanto ai «parametri principali utilizzati per programmare o addestrare i sistemi»²³.

Questo significa, ad esempio, che per comprendere se l'algoritmo che valuta le candidature ad una promozione ha tendenze discriminatorie, il legislatore confida in una verifica limitata alle categorie di dati ed alla regolazione degli *iperparametri* impiegati dagli sviluppatori per supervisionare il processo di apprendimento. Senza necessità di ricostruire ogni singolo passaggio logico del processo decisionale automatizzato.

Al fondo il decreto n. 104 cit. suggerisce alla riflessione giuslavorista applicata all'intelligenza artificiale la necessità di adottare una diversa e più compromissoria impostazione, giusta la quale, abbandonata una lettura antropomorfizzante dell'intelligenza artificiale, il *focus* si sposta da una impossibile disamina dei moventi algoritmici dei provvedimenti aziendali, alla analisi della tipologia di dati e dei processi che hanno alimentato le decisioni algoritmiche.

²⁰ Per Peruzzi 2023, 150, tanto più il sistema è autonomo tanto più si depotenzia la capacità della trasparenza «rispetto all'obiettivo di assicurare l'effettività dei diritti potenzialmente incisi dall'uso dell'IA».

²¹ Cfr. invece l'argomentato pessimismo di Gragnoli 2022 sp. 34 sgg.

²² Fra i quali anzitutto vanno annoverati, perché particolarmente evidenti, i limiti dati dalle risorse economiche processuali disponibili, su cui Gragnoli 2022.

²³ Qui, come con riguardo alle «metriche» utilizzate per misurare i parametri di accuratezza, robustezza e cybersicurezza dei sistemi, e poi gli impatti potenzialmente discriminatori delle «metriche stesse», il giurista entra a contatto con un lessico da statistica bayesiana che non gli appartiene, e con il quale deve familiarizzare. Va segnalato inoltre che l'art. 1 bis cit. impone l'informativa anche circa le «misure di controllo adottate per le decisioni automatizzate» e «gli eventuali processi di correzione e il responsabile del sistema di gestione della qualità». Questa disposizione appare criticabile considerato che l'adozione di misure di controllo e la designazione del responsabile non sono stati resi obbligatori.

Non sarà quella trasparenza algoritmica assoluta che, forse ingenuamente, anche solo pochi anni fa ancora auspicavamo sui luoghi di lavoro e che, invece, realisticamente, abbiamo compreso mai avremo. Ma è pur sempre un primo passo, mosso al fondo dalla consapevolezza che se anche avessimo accesso a ciascuno dei miliardi di dati processati da un sistema incentrato su algoritmi cd. malleabili (come lo è ad es. il motore di ricerca di *Google*), che vengono riprogrammati di continuo e che subiscono continui aggiornamenti senza alcuna influenza umana, non avremmo più trasparenza, forse avremmo persino meno trasparenza²⁴.

D'altra parte l'uomo (dunque il giudice od il consulente tecnico incaricato) non è in grado di indagare una simile mole di dati, può solo affidarsi ad altri sistemi intelligenti che lo facciano per lui.

Ed anche in tal caso, ci spiegano gli informatici, probabilmente con scarsa efficacia. Bisognerebbe infatti considerare i possibili limiti di *design* dell'algoritmo indagato, i cui codici potrebbero essere stati strutturati senza curarsi di tracciare in alcun modo gli elementi che hanno concorso alla decisione algoritmica, e la cui malleabilità strutturale potrebbe persino essere stata indotta per ottenere un intenzionale offuscamento di dati e parametri utile per «l'alleggerimento delle responsabilità» (Colaiacomo 2022, 213).

6. (segue) l'enfasi sui dati e la lotta alle algodiscriminazioni

L'enfasi sulla trasparenza evidenzia dunque un chiaro spostamento del baricentro della regolazione giuridica sulla *qualità* e sulla *governance* algoritmica dei dati.

D'altra parte l'intelligenza artificiale che auto-apprende dai dati e che elabora a partire dai dati in maniera incondizionata, proprio perché ha una consapevolezza strutturalmente limitata ai dati e finanche dei dati stessi, pur a fronte di una capacità di processo sovrumana, può incidentalmente generare informazioni errate, produrre istruzioni dannose o contenuti difforni dalla realtà²⁵.

È stato evidenziato come i sistemi di intelligenza artificiale che proseguono il loro apprendimento dopo essere stati introdotti nel mercato possono cadere preda dei cd. circuiti di *feedback* o *loop*, ossia di errori o *bias* algoritmici, dovuti alla acquisizione e poi al rinforzo di assunzioni sbagliate nel processo di apprendimento automatico.

Ciò deriva molto spesso dal fatto che i dati che l'intelligenza artificiale inconsapevolmente processa sono stati sovraccaricati dagli stessi utenti di distorsioni

²⁴ È questo un approccio per certi versi obbligato di fronte ad un'intelligenza artificiale probabilistica e connessionista che si alimenta di dati che processa per estrarre conoscenza. La disponibilità di dati infatti cresce in maniera esponenziale e difficilmente commensurabile, comunque al di fuori dal controllo umano. Secondo Floridi 2022, 25, il numero dei dispositivi digitali che interagiscono tra loro è già oggi superiore alla popolazione mondiale, con la conseguenza che la maggior parte delle comunicazioni avviene tra dispositivi intelligenti senza coinvolgimento umano. Nel senso che in queste condizioni effettivamente «l'informazione di per sé non è garanzia di tutela effettiva» cfr. Battelli (2022, 101)

²⁵ Questo del resto riportano esplicitamente le avvertenze sul sito di ChatGPT.

cognitive, pregiudizi, influenze ideologiche e di genere, opinioni e comportamenti discriminatori.

L'interpretazione algoritmica può anche generare, specie in caso di scarsa qualità dei dati, ma anche a ragione di difetti del sistema od errori nel processo di *training*, l'apofenia ossia una immotivata visione di connessioni accompagnata dall'attribuzione di una significatività spropositata²⁶.

Ricerche condotte sulla profilazione razziale involontaria hanno dimostrato che avere più dati disponibili da processare non migliora necessariamente la qualità della rappresentazione algoritmica della realtà, ed anzi può perfino esacerbare il problema, proprio per il tema del rinforzo e dei *loops*.

Lequità algoritmica è insidiata anche dal cd. formalismo algoritmico, cioè della tendenza alla neutralità, che finisce per consolidare eventuali pregiudizi o *bias* algoritmici preesistenti. E dalla cd. trappola dell'astrazione, che deriva dalla incapacità dell'algoritmo di rendersi conto del contesto in cui opera, perché gli sfugge la cornice emotivo relazionale, che, invece, in ambito lavorativo è praticamente sempre decisiva.

Questi problemi vengono poi aggravati a valle dell'elaborazione algoritmica dal cd. pregiudizio dell'automazione, che ci fa presumere che l'intelligenza artificiale ha sempre ragione²⁷.

Per tutte queste ragioni la strada da percorrere per migliorare la convivenza sui luoghi di lavoro con gli algoritmi di apprendimento autonomo è evidentemente ancora lunga ed accidentata.

Paradossalmente l'ambito per il quale è forse meno urgente l'introduzione di nuove norme giuridiche è proprio quello delle cd. algodiscriminazioni, che pure, se guardiamo alla casistica già emersa alle cronache, si dimostra particolarmente esposto.

In questo settore, infatti, il diritto del lavoro si è da tempo attrezzato per contrastare l'opacità informativa e le decisioni inspiegabili del datore di lavoro, accedendo alla nozione di discriminazione indiretta ed all'efficace meccanismo della inversione dell'onere della prova in caso di discriminazione presumibile in base a rilievi statistici²⁸.

²⁶ Il fatto che l'intelligenza artificiale possa mostrare attitudine a riconoscere schemi o connessioni tra informazioni (patterns) che non hanno invece una significativa correlazione logica, conferma che esiste evidentemente anche un tema legato alla differente qualità dei sistemi che fanno uso di questa tecnologia. Tema che meriterebbe la medesima attenzione di quello che attiene alla qualità dei dati processati. In letteratura ricorre infatti sovente l'affermazione che «un algoritmo non può che riflettere la qualità dei dati su cui è costruito» cfr. Battelli (2022, 100), rispetto alla quale è agevole convenire ma con la postilla che un algoritmo non può che riflettere, a maggior ragione, anche la qualità della sua impostazione informatica.

²⁷ La formula pregiudizio dell'automazione descrive la condizione dell'individuo che ignora o evita di cercare informazioni in contraddizione con una soluzione generata dall'intelligenza artificiale, considerandola di per sé corretta e ottimale. Cfr. anche per riferimenti Floridi 2022, 153.

²⁸ Sulla tendenza all'alleggerimento dell'onere probatorio da parte della Corte di Giustizia cfr. M. Peruzzi (2023, 159).

E si è spinto direi al limite delle proprie possibilità anche dal punto di vista sanzionatorio, con la nullità e la vasta pletora dei danni rivendicabili deputati ad innalzare la soglia di tutela in caso di discriminazione, ancorché non espressamente punitivi²⁹.

Nondimeno vi possono essere altre direttrici da percorrere per incrementare l'equità algoritmica nel diritto del lavoro.

La più efficiente, ma per nulla scontata, consisterebbe nell'intervenire già nella fase di *training* e durante la formazione dei set di dati con cui vengono addestrati gli algoritmi.

Occorrerebbe incentivare normativamente, durante la fase in cui gioca ancora un ruolo la «mano» dell'uomo, l'adozione di standard tecnici che includano processi di formazione degli algoritmi all'impiego di tecniche di *adversarial debiasing* per mitigare il rischio da ingerenza di pregiudizi «codificati», prendendo come riferimento quanto dispone l'art. 9 del Regolamento dell'Unione europea n. 679/2016, in tema di dati genetici e biometrici, origine razziale o etnica ed orientamento sessuale.

L'algoritmo andrebbe in sostanza formato anche per riuscire ad individuare e rimuovere i (propri) pregiudizi algoritmici prima ancora che i meccanismi di *loops* amplifichino la distorsione. In caso anche bilanciando con una azione algoritmica di rinforzo positivo eventuali tendenze potenzialmente discriminatorie. O ponendo dei limiti di *design* al sistema attraverso appositi macroparametri congegnati in funzione dell'obiettivo da perseguire, anche accedendo ad una tutela per quote, sia pure in ultima istanza vista l'elevata problematicità insita nello stesso rimedio.

Quel che è certo è che se la regolazione giuridica non investe nella prevenzione algoritmica, gioco forza occorrerà poi perseguire la direttrice sanzionatoria, che opera a valle del processo decisionale automatizzato.

Ma questo vuol dire condannarsi ad una improba attività di ricerca dei *bias* eventualmente responsabili, con tutte le difficoltà del caso, *in primis* con riguardo alla analisi dell'inerenza causale, in termini anche qui necessariamente probabilistici³⁰, tra le categorie di dati su cui un dato sistema intelligente è stato ad-

²⁹ Secondo Corte Cass., sez. un., 21 luglio 2021, n. 20819 la risarcibilità del danno da discriminazione non può essere ricondotta alla categoria dei danni punitivi, stante la natura eminentemente deterrente-dissuasiva dei rimedi finalizzati a garantire l'effettività della tutela antidiscriminatoria orientata alla conformità comunitaria. Nondimeno, per a dare maggiore consistenza ed effettività al danno risarcibile la giurisprudenza ammette la legittimità della liquidazione cd. ultra-compensativa del danno, tenuto adeguatamente conto che «l'atto discriminatorio è lesivo della dignità umana» cfr. Cass. 2 novembre 2021 n. 31071 e Cass. 15 dicembre 2020, n. 28646. In letteratura Biasi 2023.

³⁰ Cfr. l'art. 4 co. 1 lett. b della Proposta di Direttiva del Parlamento Europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale). Diviene sempre più evidente lo slittamento del diritto del lavoro verso un orizzonte epistemologico probabilista, Maio 2018b.

destrato ed alle quali ha continuato autonomamente ad attingere e la presunta discriminazione sul lavoro.

7. Per un diritto generalizzato alla supervisione umana delle decisioni automatizzate nel rapporto di lavoro

Nell'ottica di implementare ulteriori strumenti di tutela del lavoro oltre la trasparenza, colpisce che alcuni fondamentali diritti a contenuto eminentemente procedurale³¹ siano, allo stato, in via di adozione esclusivamente nell'ambito settoriale delle piattaforme digitali³².

Quasi che di una gestione algoritmica responsabile debbano beneficiare soltanto i lavoratori delle piattaforme³³.

Il riferimento è al Capo III della proposta di direttiva sulle piattaforme dedicato alla «*gestione algoritmica*» dei rapporti di lavoro, che, oltre agli obblighi di informazione, introduce anche principi ulteriori, a mio avviso, di necessaria generalizzazione.

Fra questi, all'art 7, la necessità di un monitoraggio umano dell'impatto dei sistemi automatizzati, e soprattutto, all'art. 8, il diritto al riesame umano di decisioni che impattano in maniera significativa sulle condizioni dei lavoratori³⁴. Cui si riconnette l'obbligo di motivazione scritta quando la decisione automatizzata conduce alla risoluzione del rapporto di lavoro o incide sulla retribuzione.

La facoltà di riesame umano delle decisioni automatizzate della piattaforma digitale ha senso solo se implica la possibilità di una loro rivalutazione e rettifica «*senza indugio*», ed anche il riconoscimento di compensazioni, anzitutto economiche ma non solo, se gli esiti della decisione algoritmica hanno violato i diritti del lavoratore.

Ebbene, quella per un diritto generalizzato alla supervisione umana di tutte le decisioni algoritmiche che incidono in maniera rilevante su aspetti fonda-

³¹ Per Tullini 2022, 52, «anziché di diritti soggettivi concettualmente individuati» si tratterebbe «piuttosto di agibilità: cioè, di declinazioni - a carattere essenzialmente procedurale - dei principi generali di trasparenza e di *accountability* relativi alla gestione algoritmica». Diversamente, per Magnani 2022 la digitalizzazione costringe ad individuare «nuovi diritti della persona» che necessariamente «si declinano in modo diverso».

³² Vd. la Proposta di Direttiva del Parlamento europeo e del Consiglio relativa al miglioramento delle condizioni di lavoro nel lavoro mediante piattaforme digitali, del 9 dicembre 2021, COM(2021)762 final, su cui fra i molti Barbieri 2021.

³³ Come già rilevato in passato Romagnoli 2018.

³⁴ Che l'art. 8 declina essenzialmente come «il diritto di ottenere una spiegazione» «per qualsiasi decisione presa o sostenuta da un sistema decisionale automatizzato che incida significativamente sulle condizioni di lavoro del lavoratore», garantendo la «possibilità di rivolgersi a una persona di contatto», che deve essere competente e formata, per «discutere e chiarire i fatti, le circostanze e i motivi di tale decisione». Si tratta di una garanzia minima che invocavamo in via generalizzata nel 2017, cfr. Maio (2018a, 1442). Nello stesso senso anche Battelli (2022, 102).

mentali del rapporto di lavoro costituisce, a mio modo di vedere, e non da oggi, la più irrinunciabile fra le battaglie per la civilizzazione della società digitale³⁵.

Perché il diritto del lavoro, al di là dello specifico delle piattaforme digitali, resta pur sempre un diritto presidiato da norme espressive di un bilanciamento valoriale complesso, che non di rado si riflette in clausole generali aperte all'interpretazione, le quali sono state pensate per venire rimesse alla valutazione umana, e non certo a quella algoritmica.

8. Per un divieto generalizzato di impiego delle tecnologie induttive e manipolatorie nei luoghi di lavoro (ancora su lavoro e gamificazione)

Tra le previsioni della proposta di direttiva sulle piattaforme digitali che meriterebbero una applicazione generalizzata alla gestione algoritmica *tout court* di ogni rapporto di lavoro c'è anche il divieto di trattamento dei dati «personali relativi allo stato emotivo o psicologico del lavoratore» (art. 6, co. 5, lett. a).

Più in generale, lascia presagire scenari inquietanti il possibile impiego nei luoghi di lavoro di sistemi automatizzati che utilizzano l'intelligenza artificiale per la cd. deduzione delle emozioni (si pensi all'utilizzo nella selezione di un candidato per una promozione lavorativa).

Questa preoccupazione trova conferma nella già citata proposta di Regolamento sull'intelligenza artificiale dove i sistemi per il riconoscimento delle emozioni sono oggetto di una specifica attenzione ed originano un apposito obbligo di trasparenza (cfr. artt. 1 lett. c) e 52 co. 2)³⁶.

In proposito occorre considerare che nella prospettiva del cd. *Artificial Intelligence Act* la nozione di sistema intelligente deputato a dedurre le emozioni, include anche la capacità di decifrare le intenzioni dell'interlocutore umano³⁷.

I processi automatizzati di deduzione delle emozioni e delle intenzioni incidono direttamente sulla dignità della persona che lavora, in quanto vanno potenzialmente ad intaccare quello che potremmo definire l'*habeas corpus* digitale del lavoratore, dal momento che lo privano del diritto di mantenere spazi di privacy e indeducibilità.

La dignità è direttamente connessa anche con il divieto per i «sistemi decisionali e di monitoraggio automatizzati» di mettere «in alcun modo» «indebitamente sotto pressione i lavoratori»; divieto che è pure contenuto nella proposta di direttiva sulle piattaforme digitali (art. 7), mentre sembrerebbe meritevole di una più generale valorizzazione.

³⁵ Maio 2018a. Sul principio del «controllo umano» insiste anche l'Accordo Quadro delle Parti Sociali Europee sulla digitalizzazione del giugno 2020. Favorevoli alla revisione umana anche Oddenino 2020 e Peruzzi 2023.

³⁶ Cfr. Incutti 2022.

³⁷ Per l'art. 3 n. 34 la formula di «sistema di riconoscimento delle emozioni» descrive un sistema di intelligenza artificiale «finalizzato all'identificazione o alla deduzione di emozioni o intenzioni di persone fisiche sulla base dei loro dati biometrici».

La formula dell'indebita pressione – evidentemente generica, ma certo più efficace descrittivamente di un anonimo richiamo allo *stress* lavoro correlato – coglie infatti una peculiare minaccia all'equilibrio psico-fisico del lavoratore che può conseguire dal *modus operandi* dell'intelligenza artificiale³⁸.

L'esperienza di questi anni ha infatti dimostrato come l'intelligenza algoritmica sia in grado di esercitare anche in ambito lavorativo un vero e proprio potere induttivo, mediante raccomandazioni continue e pervasive, dosando premialità e minaccia sanzionatoria, con l'obiettivo di conformare progressivamente l'azione del proprio interlocutore. Il potenziale profilattivo e predittivo dell'intelligenza artificiale può, in sostanza, venire posto al servizio di una incessante spinta gentile, che interagisce con l'autonomia dei lavoratori per modellarne le scelte, suggerendo o reclamando conformazione³⁹. A dispetto di una formale condizione di autodeterminazione del lavoratore.

La «pressione» esercitata dall'intelligenza artificiale, che in una misura per così dire fisiologica sarebbe ancora da riconnettere al vincolo stesso di subordinazione, diviene però indebita quando si traduce in una costante tensione manipolatoria, che si avvale della carenza di comprensione delle dinamiche che muovono la valutazione algoritmica o della incapacità di mettere in discussione i risultati della decisione algoritmica.

Non a caso, nella gestione del personale assume sempre più un ruolo determinante la cd. gamificazione del lavoro, ossia la tendenza all'impiego sistematico dei meccanismi tipici del gioco (e, in particolare, dei videogiochi), come l'assegnazione di punti, l'inserimento in *rating*, il superamento di livelli, il conseguimento di premi e facoltà (Maio 2022, sp. 49-51).

L'intelligenza artificiale può così consentire di gestire in tempo reale il ricorso pianificato ai meccanismi di ricompensa ed ai cd. rinforzi positivi, che le neuroscienze hanno chiarito costituiscono leve più efficaci della minaccia di sanzioni, perché, attivando emozioni positive che coinvolgono il piacere, inducono comportamenti appetitivi che generano adempimento per motivazione.

Potremmo forse chiarire meglio questa tendenza, proprio riprendendo l'esempio di poc'anzi. Per come mi pare di intenderne gli sviluppi, l'intelligenza artificiale sui luoghi di lavoro non verrà impiegata per scovare la finta morbilità, servirà piuttosto per indurre il lavoratore a fare sempre meno ricorso ai propri diritti di malattia.

E questo, mi sembra, dovrebbe preoccuparci molto di più, e perciò indurci a ragionare anche di un divieto generalizzato di impiego dell'intelligenza artificiale con finalità induttivo-manipolatorie della libera determinazione di chi lavora.

³⁸ Ed infatti il divieto ed il conseguente limite di impiego si estendono anche ad altre modalità operative del sistema intelligente che si rivelino tali da mettere «altrimenti a rischio la salute fisica e mentale dei lavoratori delle piattaforme digitali».

³⁹ Abbiamo già avuto modo di rilevare questa tendenza in passato sia con riguardo alla robotica intelligente (Maio 2018a, 1441), sia con riferimento alle piattaforme digitali (Maio 2019, 591).

Perché è banale a dirsi, ma la civilizzazione del lavoro digitale richiede non soltanto un complessivo non arretramento delle tutele, ma anche il persistente rispetto dell'autonomia della persona che lavora.

Ed è invece scontato vaticinare che una ingegnerizzazione incontrollata dei luoghi di lavoro condurrà inevitabilmente ad un'organizzazione del lavoro in cui l'intelligenza artificiale è di casa (Floridi 2022, 58), mentre la persona umana sarà sempre più spaesata, indifesa, coercibile.

La sfida posta oggi dalla depersonalizzazione del lavoro digitale è, dunque, nuova ed antica allo stesso tempo. Perché al fondo si tratta, ancora una volta, di riuscire a partorire una equilibrata convivenza tra le esigenze contrapposte del capitalismo digitale e della persona che lavora.

Il complesso della regolamentazione sull'intelligenza artificiale antropocentrica in corso di elaborazione in sede euro-unitaria in questo senso fa ben sperare, se di lì il legislatore muoverà per declinare poi uno specifico statuto protettivo per chi lavora relazionandosi con forme di intelligenza artificiale autonome e capaci di apprendimento nel senso dappprincipio chiarito.

A patto dunque di non accontentarsi di soluzioni compromissorie di taglio essenzialmente consumeristico, che suonerebbero per il giuslavorista inevitabilmente al ribasso, quando non inammissibili.

Perché, è sempre utile ricordare, «il pieno sviluppo della persona umana» resta il fine della nostra Costituzione lavorista e personalista (Maio 2020b), e la «dignità umana» l'argine «invalicabile» in ogni luogo di lavoro⁴⁰ anche rispetto alle manifestazioni algoritmiche dell'iniziativa economica.

Riferimenti bibliografici

- Alpa, G. 2022. "Quale modello normativo europeo per l'intelligenza artificiale?" In *Intelligenza artificiale e diritti della persona*, a cura di D. Buzzelli, M. Palazzo, 17 sgg. Pisa: Pisa University Press.
- Ans, B., Rousset, S., French, R. M., Musca, S. 2004. "Self-refreshing memory in artificial neural networks: learning temporal sequences without catastrophic forgetting." *Connection Science* 16: 71-99. <<https://doi.org/10.1080/09540090412331271199>>
- Barbieri, M. 2021. "Prime osservazioni sulla proposta di direttiva per il miglioramento delle condizioni di lavoro nel lavoro con piattaforma." *Labour Law Issues* 2: 3 sgg.

⁴⁰ Rodotà 2010, 7-8: «Proprio gli articoli 36 e 41 della Costituzione forniscono una indicazione preziosa per affrontare la questione del rapporto tra libertà e dignità, partendo da una indispensabile distinzione. L'articolo 41 indica nella dignità un limite invalicabile per l'iniziativa economica privata; l'articolo 36 indica il criterio per la costruzione della dignità e per l'individuazione del soggetto al quale spetta questo potere... Ricordo ancora che quest'ultimo articolo parla di «esistenza libera e dignitosa». La costruzione dell'homo *dignus* non può essere effettuata all'esterno della persona, ha davvero il suo fondamento in interiore *homine*. La dignità non è indeterminata, ma trova nella persona il luogo della sua determinazione, tuttavia non per custodire un'essenza, bensì per mettere ciascuno nella condizione di determinare liberamente il proprio progetto di vita».

- Battelli, E. 2022. "Necessità di un umanesimo tecnologico: sistemi di intelligenza artificiale e diritti della persona." In *Intelligenza artificiale e diritti della persona*, a cura di D. Buzzelli, M. Palazzo, 101 sgg. Pisa: Pisa University Press.
- Biasi, M. 2023. *Studio sulla polifunzionalità del risarcimento del danno nel diritto del lavoro: compensazione, sanzione, deterrenza*. Milano: Giuffrè.
- Buzzelli, D., Palazzo, M. (a cura di). 2022. *Intelligenza artificiale e diritti della persona*. Pisa: Pisa University Press.
- Carinci, M. T., Giudici, S., Perri, P. 2023. "Obblighi di informazione e sistemi decisionali e di monitoraggio automatizzati (art. 1-bis «Decreto Trasparenza»): quali forme di controllo per i poteri datoriali algoritmici?" *Labor* 1: 7 sgg.
- Colaiacono, G. 2022. "Intelligenza artificiale e dignità del lavoratore: la necessità di un approccio antropocentrico." In *Intelligenza artificiale e diritti della persona*, a cura di D. Buzzelli, M. Palazzo, 205 sgg. Pisa: Pisa University Press.
- Costanza, M. 2019. "L'intelligenza artificiale e gli stilemi della responsabilità civile." *Giurisprudenza Italiana*, 686 sgg.
- Cristianini, N. 2023. *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*. Bologna: Il Mulino.
- D'Aloia, A. (a cura di). 2020. *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*. Milano: Franco Angeli.
- Del Punta, R. 2018. "Un diritto per il lavoro 4.0." In *Il lavoro 4.0: la Quarta Rivoluzione industriale e le trasformazioni delle attività lavorative*, a cura di A. Cipriani, A. Gramolati, G. Mari, 225 sgg. Firenze: Firenze University Press.
- Faioli, M. 2023. "Giustizia contrattuale, tecnologia avanzata e reticenza informativa del datore di lavoro. Sull'imbarazzante «truismo» del decreto trasparenza." *Diritto delle Relazioni Industriali* 1, 1: 45 sgg.
- Floridi, L. 2022. *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*. Milano: Raffaello Cortina.
- Garofalo, D., Tiraboschi, M. 2023. "L'impatto sulla disciplina del rapporto di lavoro del d.lgs. n. 104/2022 (emanato in attuazione della direttiva UE n. 2019/1152)." In *Trasparenza e attività di cura nei contratti di lavoro. Commentario ai decreti legislativi n. 104 e n. 105 del 2022*, a cura di D. Garofalo, M. Tiraboschi, V. Fili, A. Trojsi, 10 sgg. Bergamo: ADAPT University Press.
- Garofalo, D., Tiraboschi, M., Fili, V., Trojsi, A. (a cura di). 2023. *Trasparenza e attività di cura nei contratti di lavoro. Commentario ai decreti legislativi n. 104 e n. 105 del 2022*. Bergamo: ADAPT University Press.
- Gragnoli, E. 2020. "Il rapporto di lavoro, le decisioni automatiche e i loro limiti. La regolazione di fenomeni incomprensibili." In *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, a cura di A. D'Aloia, 461 sgg. Milano: Franco Angeli.
- Gragnoli, E. 2022. "Il potere di controllo, le risorse digitali e gli algoritmi." In *Tecnologie digitali, poteri datoriali e diritti dei lavoratori*, a cura di A. Bellavista, R. Santucci, 27 sgg. Torino: Giappichelli.
- Incutti, E. M. 2022. "Sistemi di riconoscimento delle emozioni e ruolo dell'autonomia privata: linee evolutive di un umanesimo digitale." In *Intelligenza artificiale e diritti della persona*, a cura di D. Buzzelli, M. Palazzo, 151 sgg. Pisa: Pisa University Press.
- Magnani, M. 2022. "La proposta di direttiva sul lavoro mediante piattaforme digitali." *Audizione presso la XI Commissione (Lavoro pubblico e privato) della Camera dei Deputati, XVIII Legislazione, esame della Proposta di direttiva del Parlamento europeo e del Consiglio, relativa al miglioramento delle condizioni di lavoro nel lavoro mediante piattaforme digitali*, 21 aprile 2022. In *Bollettinoadapt.it*.

- Maio, V. 2018a. "Il diritto del lavoro e le nuove sfide della rivoluzione robotica." *Argomenti di diritto del lavoro*, 1415 sgg.
- Maio, V. 2018b. "La perdita di chances nel rapporto di lavoro e le recenti evoluzioni in tema di responsabilità civile." In *Argomenti di diritto del lavoro*, 101 sgg.
- Maio, V. 2019. "Il lavoro per le piattaforme digitali tra qualificazione del rapporto e tutele." *Argomenti di diritto del lavoro*, 591 sgg.
- Maio, V. 2020a. "Diritto del lavoro e potenziamento umano. I dilemmi del lavoratore aumentato." *Giornale di diritto e relazioni industriali*, 513 sgg.
- Maio, V. 2020b. "Il Titolo primo dello Statuto e il principio personalista." In *Mezzo secolo dallo Statuto dei lavoratori. Politiche del diritto e cultura giuridica*, a cura di M. Rusciano, L. Gaeta, L. Zoppoli. Quaderni della Rivista Diritti Lavori Mercati vol. 8, 165 sgg. Napoli: Editoriale Scientifica.
- Maio, V. 2022. "Diritto del lavoro e metaverso. Se il lavoro non è un (video) gioco." *Labour Law Issues* 8: 42 sgg.
- Marazza, M. e D'Aversa, F. 2022. "Dialoghi sulla fattispecie dei «sistemi decisionali o di monitoraggio automatizzati» nel rapporto di lavoro (a partire dal decreto trasparenza)." *Giustiziacivile.com*, 8 novembre 2022.
- Mazlish, B. 2004. *Civilization and Its Contents*. Stanford: Stanford University Press.
- Oddenino, A. 2020. "Decisioni algoritmiche e prospettive internazionali di valorizzazione dell'intervento umano." *Diritto pubblico comparato ed europeo online* 1: 203 sgg.
- Pasquale, F. 2016. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard: Harvard University Press.
- Peruzzi, M. 2023. *Intelligenza artificiale e lavoro. Uno studio su poteri datoriali e tecniche di tutela*. Torino: Giappichelli.
- Pietropaolo, S. 2022. "La maschera di silicio. Un requiem per il diritto dell'era digitale." In *Intelligenza artificiale e diritti della persona*, a cura di D. Buzzelli, M. Palazzo, 39 sgg. Pisa: Pisa University Press.
- Proia, G. 2022a. "Le novità, sistemiche e di dettaglio, del decreto trasparenza." *Massimario di giurisprudenza del lavoro* 3: 571 sgg.
- Proia, G. 2022b. "Trasparenza, prevedibilità e poteri dell'impresa." *Labor* 6: 641 sgg.
- Rodotà, S. 2010. *Antropologia dell'homo dignus*. Lezione tenuta il 6 ottobre 2010 in occasione del conferimento della Laurea honoris causa dall'Università di Macerata. Open access.
- Romagnoli, U. 2018. "Se l'amore per la specie fa perdere di vista il genere (a proposito del caso Foodora)." *Diritti Lavori Mercati* 2: 193 sgg.
- Santoro Passarelli, G. 2022. "Dignità del lavoratore e libertà di iniziativa economica, rivoluzione industriale 4.0 e rapporto di lavoro dei riders." *Labor* 1: 5 sgg.
- Schiavone, E. C. 2023. "Gli obblighi informativi in caso di sistemi decisionali o di monitoraggio automatizzati." In *Trasparenza e attività di cura nei contratti di lavoro. Commentario ai decreti legislativi n. 104 e n. 105 del 2022*, a cura di D. Garofalo, M. Tiraboschi, V. Fili, A. Trojsi, 211 sgg. Bergamo: ADAPT University Press.
- Tullini, P. 2022. "La direttiva piattaforme e i diritti del lavoro digitale." *Labour Law Issues* 8: 52.
- Zilli, A. 2023. "La via italiana per condizioni di lavoro trasparenti e prevedibili." *Diritto delle relazioni industriali* 1: 25 sgg.