



Corso di laurea di Giurisprudenza

Cattedra di Diritto ed Etica delle Tecnologie Emergenti

# Intelligenza Artificiale e lavoro: Tra Etica e Diritto

Chiar.ma Prof.ssa Laura Pallazzani

---

RELATRICE

Chiar.ma Prof.ssa Alessia Farano

---

CORRELATRICE

Elisa

Consiglio

Matr.178053

---

CANDIDATA

Anno Accademico 2025/26

*Alle meravigliose donne della mia vita, mia mamma Marika, mia zia Lalla, mia nonna MariaPia e le mie zie Cettina e Pia. Siete state: maestre di vita, ali con cui volare, fate turchine, lucciole nelle notti oscure, guerriere senza paura, parole intrinseche di forza, la magia di cui è circondata la mia vita, il mio scudo di luce contro il male del mondo, la nave che mi ha portato ad Itaca e verso l'Isola che non c'è.*

# INDICE

<u>INTRODUZIONE.....</u>	<u>7</u>
<u>CAPITOLO 1.....</u>	<u>9</u>
<u>FONDAMENTI TEORICI, ETICI E NORMATIVI DELL'INTELLIGENZA ARTIFICIALE ...</u>	<u>9</u>
<u>1. L'INTELLIGENZA ARTIFICIALE.....</u>	<u>9</u>
1.1 DEFINIZIONE TECNICA E CONCETTUALE.....	9
1.2 L'IA COME PARADIGMA CULTURALE SECONDO GRASSI .....	12
<u>2. LA QUARTA RIVOLUZIONE INDUSTRIALE: EVOLUZIONE STORICA E AMBITI DI APPLICAZIONE.....</u>	<u>13</u>
2.1 UNA RIVOLUZIONE TECNOLOGICA E ANTROPOLOGICA .....	13
2.2 LAVORO E MUTAZIONE DELLE COMPETENZE.....	13
2.3 ECOSISTEMI DELL'INNOVAZIONE E TRANSDISCIPLINARITÀ .....	14
2.4 AI NELLA SANITÀ: LA RIVOLUZIONE SILENZIOSA.....	15
2.5 DIFESA E SICUREZZA STRATEGICA.....	16
2.6 INGEGNERIA E SOSTENIBILITÀ AMBIENTALE .....	17
2.7 AGRICOLTURA INTELLIGENTE E TECNOLOGIE SOSTENIBILI .....	17
2.8 LOGISTICA E AUTOMAZIONE DELLA FILIERA PRODUTTIVA .....	17
2.9 FINANZA, BANCHE E MONETA DIGITALE.....	18
2.10 ISTRUZIONE E AMBIENTI DI APPRENDIMENTO DIGITALI .....	18
2.11 AEROSPAZIO E COMPETIZIONE STRATEGICA GLOBALE.....	18
2.12 GOVERNANCE PUBBLICA E REGOLAZIONE ALGORITMICA.....	19
<u>3. IL DIBATTITO ETICO .....</u>	<u>19</u>
3.1 L'IMMAGINARIO COLLETTIVO TRA UTOPIA E DISTOPIA .....	19
3.2 COSCIENZA ARTIFICIALE E DIGNITÀ UMANA: VERSO UN'ETICA DELLA REGOLAZIONE.....	20
<u>4. L'ALGORETICA: UNA NUOVA ETICA PER I SISTEMI INTELLIGENTI .....</u>	<u>22</u>
4.1 SPIEGABILITÀ E DIGNITÀ UMANA.....	24
4.2 NON DISCRIMINAZIONE ALGORITMICA .....	26
4.3 IL CONTROLLO UMANO SIGNIFICATIVO.....	29
4.4 LA RESPONSABILITÀ NELL'ERA DELL'IA.....	30
4.5 UNA TECNOLOGIA DA ORIENTARE: IL COMPITO DELL'ETICA TRA SFIDE E RESPONSABILITÀ... 32	

<u>5. ETICA E DIRITTO IN DIALOGO .....</u>	<u>33</u>
5.1 IL PROGETTO <i>ETHICALLY ALIGNED DESIGN</i> DELL'IEEE .....	34
5.2 IL CONTRIBUTO DELL'HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE.....	36
5.3 OECD AI PRINCIPLES (2019).....	38
5.4 L'APPROCCIO ETICO DELL'UNESCO ALL'INTELLIGENZA ARTIFICIALE.....	39
<u>6. DALL'ETICA AL DIRITTO: LA SVOLTA NORMATIVA DELL'UNIONE EUROPEA .....</u>	<u>41</u>
6.1 AI ACT.....	42
6.2 ARCHITETTURA NORMATIVA DELL'AI ACT E APPROCCIO BASATO SUL RISCHIO .....	44
6.3 LE TENSIONI INTERNE DELL'AI ACT: TRA AMBIZIONE REGOLATIVA E FRAGILITÀ STRUTTURALE .....	46
<u>7. LE INIZIATIVE LEGISLATIVE ITALIANE IN MATERIA DI INTELLIGENZA ARTIFICIALE: IL D.D.L. N. 1146-B.....</u>	<u>49</u>
<u>7.1 L'APPLICAZIONE SETTORIALE DELL'INTELLIGENZA ARTIFICIALE NEL DISEGNO DI LEGGE N. 1146-B.....</u>	<u>52</u>
<u>CAPITOLO 2.....</u>	<u>55</u>
<u>LAVORO, INTELLIGENZA ARTIFICIALE E TUTELA DEI DIRITTI FONDAMENTALI ..</u>	<u>55</u>
1. TRASFORMAZIONI DEL LAVORO NELL'ERA DELL'INTELLIGENZA ARTIFICIALE.....	55
1.2 IL PARADIGMA 4.0: UNA TRASFORMAZIONE TECNOLOGICA, SOCIALE E SIMBOLICA .....	56
1.3 I SETTORI ECONOMICI PIÙ INCLINI ALL'ADOZIONE DELL'IA.....	58
<u>2. LAVORO, DISABILITÀ E INTELLIGENZA ARTIFICIALE: OPPORTUNITÀ DI INCLUSIONE E RISCHI DI ESCLUSIONE SOCIALE.....</u>	<u>60</u>
2.1 INCLUSIONE LAVORATIVA DELLE PERSONE CON DISABILITÀ.....	60
2.2 IL DIGITAL DIVIDE COME BARRIERA INVISIBILE ALL'INCLUSIONE .....	63
<u>3. DISUGUAGLIANZE DIGITALI: I BIAS ALGORITMICI NEL MERCATO DEL LAVORO .....</u>	<u>66</u>
3.1 BIAS ALGORITMICI E DISCRIMINAZIONI NEI PROCESSI DECISIONALI.....	66
3.2 COME SI GENERANO I BIAS ALGORITMICI.....	67
3.3 DATI STORICI E ORIGINE DELLE DISTORSIONI ALGORITMICHE .....	69
3.4 IL CASO AMAZON .....	70
3.5 BIAS ALGORITMICI NEL SETTORE SANITARIO .....	72

3.6 VERSO SOLUZIONI EFFICACI: PREVENZIONE, CONTROLLO E RESPONSABILITÀ NELLE DISCRIMINAZIONI ALGORITMICHE.....	74
<u>4. TECNOLOGIE ALGORITMICHE E TRASFORMAZIONE DEL RAPPORTO DI LAVORO</u> .....	<u>77</u>
4.1 DIGITAL TAYLORISM E RIDEFINIZIONE DEL RUOLO MANAGERIALE .....	81
4.2 AUTOMAZIONE DECISIONALE E VALUTAZIONE DELLA PERFORMANCE .....	83
4.3 IMPATTO SULL' AUTONOMIA E SULLA DIGNITÀ DEL LAVORATORE .....	86
4.4 PRIVACY E SORVEGLIANZA ALGORITMICA NEL CONTESTO LAVORATIVO .....	88
<u>5. CONTROLLO ALGORITMICO E AUTONOMIA DECISIONALE DEL LAVORATORE.</u>	<u>92</u>
<u>6. LO “SMART WORKING”</u> .....	<u>94</u>
6.1 INQUADRAMENTO NORMATIVO E SFIDE DI TUTELA.....	96
<u>7. NORMATIVA NAZIONALE ED EUROPEA APPLICABILE</u> .....	<u>99</u>
5.2 IL QUADRO NORMATIVO NAZIONALE.....	102
<u>CAPITOLO 3</u> .....	<u>106</u>
<u>UN'INTELLIGENZA ARTIFICIALE PER L'INCLUSIONE: IL PROGETTO “ENGAGE” A TILBURG</u> .....	<u>106</u>
<u>1. CONTESTO E OBIETTIVO DEL PROGETTO</u> .....	<u>107</u>
<u>2. LA METODOLOGIA ADOTTATA: IL DESIGN THINKING COME STRUMENTO DI INCLUSIONE.</u> .....	<u>108</u>
1. EMPATIZZARE: ASCOLTARE PRIMA DI PROGETTARE.....	109
2. DEFINIRE: DARE UN NOME ALLA COMPLESSITÀ, TRASFORMARE IL PROBLEMA IN MISSIONE 111	
3. IDEARE, DALL'ASCOLTO ALL'INNOVAZIONE: CO-PROGETTARE SOLUZIONI DIGITALI PER L'INCLUSIONE.....	112
4. PROTOTIPARE, DARE FORMA ALL'IDEA: COSTRUIRE UNO STRUMENTO ACCESSIBILE, UMANO, FUNZIONALE.....	113
5. TESTARE, IL PROTOTIPO INCONTRA LA COMUNITÀ: ASCOLTARE, APPRENDERE, MIGLIORARE 115	
<u>3. LA PIATTAFORMA: STRUTTURA, FUNZIONALITÀ E FINALITÀ</u> .....	<u>117</u>

<u>4. VERSO UNA GOVERNANCE ETICA E PARTECIPATIVA DELLA PIATTAFORMA ....</u>	<u>119</u>
<u>5. IL GRUPPO CONTRO AL RESKILLING .....</u>	<u>121</u>
<u>6. IMPLICAZIONI ETICO-SOCIALI DEL PROGETTO ENGAGE .....</u>	<u>123</u>
<u>7. CONCLUSIONE: UN MODELLO REPLICABILE DI INNOVAZIONE SOCIALE E GIUSTIZIA DIGITALE.....</u>	<u>125</u>
<u>CONCLUSIONI.....</u>	<u>129</u>
<u>BIBLIOGRAFIA .....</u>	<u>131</u>
<u>SITOGRAFIA .....</u>	<u>136</u>

## INTRODUZIONE

L'intelligenza artificiale rappresenta una delle trasformazioni più incisive e pervasive della contemporaneità. Non si tratta soltanto di un insieme di tecniche informatiche, ma di un fenomeno complesso che attraversa tutti i settori della vita sociale ed economica, ridisegnando i rapporti di lavoro, i modelli di produzione e persino le categorie giuridiche ed etiche con cui interpretiamo il reale. La sua diffusione sollecita domande che vanno ben oltre il campo tecnologico: riguarda la dignità delle persone, la giustizia sociale, il futuro del lavoro e la possibilità di una convivenza basata sulla fiducia.

Questa tesi intende analizzare l'impatto dell'intelligenza artificiale con particolare attenzione al mondo del lavoro, indagando le trasformazioni che essa produce sul piano delle competenze richieste, delle tutele giuridiche e delle relazioni sociali. L'obiettivo è mostrare come l'IA non sia soltanto una tecnologia, ma una sfida antropologica, giuridica ed etica che investe i fondamenti stessi della cittadinanza democratica.

Il percorso si articola in più momenti. In primo luogo, verranno ricostruite le basi teoriche, informatiche ed etiche dell'intelligenza artificiale, così da delineare il quadro di riferimento entro cui si collocano le questioni giuridiche. In seguito, l'analisi si concentrerà sugli aspetti normativi, con particolare riferimento al contesto europeo e all'AI Act, che segna il primo tentativo organico di disciplinare i sistemi di IA attraverso un approccio basato sul rischio. Tale regolamento, pur non esente da criticità, mostra come i principi etici possano essere tradotti in vincoli giuridici, trasformando valori come trasparenza, non discriminazione e controllo umano in obblighi concreti.

Una parte significativa sarà dedicata al tema dell'inclusione e dell'esclusione. Il digital divide, nelle sue dimensioni tecnologiche, culturali e simboliche, rappresenta una delle principali sfide connesse all'IA: se da un lato essa promette maggiore accessibilità, dall'altro rischia di ampliare le disuguaglianze, soprattutto per le persone con disabilità o appartenenti a categorie fragili. Verrà messo in evidenza come la tecnologia non sia mai neutra, ma rifletta modelli impliciti di normalità che possono marginalizzare o, al contrario, valorizzare la diversità.

Accanto all'analisi teorica e normativa, sarà richiamata anche la dimensione concreta, attraverso l'esame del progetto *Engage* sviluppato a Tilburg. Questa esperienza dimostra come l'IA possa essere orientata a fini inclusivi, favorendo l'incontro tra domanda e offerta di lavoro e promuovendo opportunità per i soggetti più vulnerabili. Essa costituisce un esempio di come la tecnologia, se guidata da principi etici e da un disegno giuridico adeguato, possa rafforzare la coesione sociale invece di eroderla.

A concludere, la riflessione sarà arricchita dal richiamo a categorie filosofiche che offrono una chiave interpretativa più profonda, come vulnerabilità e fiducia. La vulnerabilità, lungi dall'essere mera debolezza, si rivela fondamento della relazione sociale e presupposto della fiducia reciproca: l'IA diventa così banco di prova della capacità collettiva di custodire questa condizione umana costitutiva.

In tal modo, la tesi si propone di offrire una visione complessiva e critica del rapporto tra intelligenza artificiale, diritto ed etica, evidenziando la necessità di governare la trasformazione tecnologica non come un processo neutrale, ma come un'occasione per rafforzare i valori fondamentali della dignità, della giustizia e della convivenza sociale.

# CAPITOLO 1

## Fondamenti teorici, etici e normativi dell'Intelligenza artificiale

L'intelligenza artificiale è ormai parte integrante delle nostre istituzioni e della società contemporanea, e sta radicalmente modificando la nostra quotidianità. Non è più una promessa tecnologica, ma una presenza operativa costante, che prende decisioni, orienta scelte e suggerisce nuovi percorsi.

Il primo capitolo di questa tesi tratta una questione preliminare ma essenziale: come costruire una cornice etico-giuridica capace di governare l'intelligenza artificiale senza esserne governata. Enunciare principi come dignità umana, trasparenza o non discriminazione non è sufficiente: essi non possono restare parole vuote, ma devono tradursi in obblighi giuridici, responsabilità chiare e garanzie concrete per i cittadini. L'indagine muove dai contributi valoriali proposti da comitati etici ed organismi internazionali, per poi analizzare il progressivo passaggio verso la normazione giuridica

In un'epoca in cui il potere si esprime sempre più attraverso il calcolo, servono strumenti che restituiscano voce, controllo e tutela alla persona. Questo è il cuore della sfida: non solo comprendere l'algoritmo, ma ridefinire i confini tra uomo e tecnica.

## 1. L'intelligenza artificiale

### 1.1 Definizione tecnica e concettuale

L'intelligenza artificiale (AI) è un ramo dell'informatica che si occupa di ideare e sviluppare algoritmi in grado di imitare e riprodurre alcune capacità cognitive e mentali, considerate per lungo tempo prerogative esclusivamente umane, come il ragionamento, l'apprendimento e il processo decisionale.<sup>1</sup> L'AI è in grado di creare software capaci di analizzare grandi quantità di dati (Big Data)<sup>2</sup>, riconoscere pattern complessi e migliorare

---

<sup>1</sup> Commissione Europea, *Cosa si intende per intelligenza artificiale (IA)*, in “Digital Strategy”, <https://digital-strategy.ec.europa.eu/it/policies/artificial-intelligence>, ultimo accesso: luglio 2025

<sup>2</sup> Osservatorio Big Data & Business Analytics, *Big Data e Business Analytics*, Politecnico di Milano, <https://www.osservatori.net/it/ricerche/osservatori-attivi/big-data-business-analytics>, ultimo accesso: luglio 2025

le proprie prestazioni nel tempo attraverso tecniche di apprendimento automatico (*machine learning*) e apprendimento profondo (*deep learning*).<sup>3</sup>

L'intelligenza artificiale, dal punto di vista informatico, si sviluppa attraverso diversi algoritmi, capaci di apprendere dai dati a sua disposizione e migliorare progressivamente le proprie prestazioni, senza che sia necessario alcun intervento umano. Il *machine learning* rappresenta quell'insieme di tecniche computazionali che permette a un sistema di riconoscere schemi e regolarità all'interno di grandi volumi di dati, di formulare inferenze e di aggiornare autonomamente i propri modelli sulla base dell'esperienza.

All'interno di questo scenario evolutivo si colloca il *deep learning*, una branca del machine learning che si distingue per l'utilizzo di **reti neurali artificiali profonde**, composte da molteplici strati gerarchici di unità computazionali. Queste reti, ispirate nella struttura al funzionamento del cervello umano, consentono l'elaborazione di dati complessi e non strutturati come -immagini, linguaggio naturale o segnali audio- senza la necessità di intervento umano nella fase di estrazione delle caratteristiche salienti. La loro efficacia risiede nella sinergia tra tre fattori chiave: la disponibilità di enormi quantità di dati (*big data*), l'accesso a elevate risorse computazionali (come *GPU* e architetture distribuite) e l'impiego di algoritmi sofisticati, tra cui la retropropagazione dell'errore (*backpropagation*).<sup>4</sup>

Quest'ultima rappresenta un meccanismo fondamentale attraverso cui la rete neurale apprende: consiste nel calcolare l'errore tra l'output generato e il risultato atteso, e nel propagarlo all'indietro lungo la rete, aggiornando progressivamente i pesi delle connessioni sinaptiche in modo da ridurre l'errore complessivo. È proprio grazie a questo processo iterativo, che consente un costante affinamento delle previsioni, che le reti profonde riescono a migliorare autonomamente le proprie prestazioni nel tempo. L'adozione di queste rivoluzionarie tecnologie ha segnato un salto qualitativo

---

<sup>3</sup> IBM, *Cos'è il deep learning?* <https://www.ibm.com/it-it/topics/deep-learning>, ultimo accesso: luglio 2025.

<sup>4</sup> TJOA, M., *A Step by Step Backpropagation Example*, in "Towards Data Science", <https://towardsdatascience.com/a-step-by-step-backpropagation-example-915fcf6cbee6>, ultimo accesso: luglio 2025.

significativo nella capacità delle macchine di replicare compiti cognitivi tradizionalmente riservati all'essere umano.

Secondo il professore Gianluca Cotaldi, l'utilizzo del termine "*intelligenza*" per descrivere questi sistemi risulta, in molti casi, eccessivamente enfatico. Non si tratta infatti di entità dotate di volontà propria o di una reale autonomia decisionale nel senso umano del termine. Più correttamente, siamo di fronte a strumenti computazionali avanzati, in grado di trarre conclusioni statisticamente fondate sulla base dell'elaborazione di esperienze pregresse e l'individuazione di schemi e correlazioni spesso invisibili ai metodi informatici tradizionali. Il grande aspetto innovativo risiede dunque nella capacità di connettere elementi profondamente eterogenei, producendo risultati predittivi o prescrittivi, che simulano senza replicare totalmente i processi decisionali umani.<sup>5</sup>

Come osserva Marco Somalvico<sup>6</sup>, tra i pionieri italiani dell'intelligenza artificiale, questa disciplina può essere considerata anche una branca dell'ingegneria, poiché orientata non solo comprensione dei processi cognitivi, ma soprattutto alla costruzione di macchine e sistemi capaci di risolvere problemi complessi attraverso comportamenti che simulano l'intelligenza umana. L'AI, in questa prospettiva, non si limita a un'analisi teorica della mente o del pensiero, ma si configura come una vera e propria disciplina progettuale, in grado di realizzare strumenti concreti ed efficienti, capaci di potenziare le prestazioni operative delle macchine e di affrontare sempre più compiti complessi in maniera quasi totalmente autonoma.

Nel corso del tempo, l'intelligenza artificiale ha seguito percorsi diversi, talvolta divergenti, nella sua evoluzione. Da un lato, ci sono gli approcci cosiddetti *logico-simbolici*, basati sulla rappresentazione chiara ed esplicita della conoscenza, come se la mente potesse essere tradotta in un linguaggio matematico perfettamente codificabile. Dall'altro lato, invece, sono emerse le tecniche *sub-simboliche*, come le *reti neurali*, che

---

<sup>5</sup> CONTALDI, G., *Intelligenza artificiale e dati personali*, in "Ordine internazionale e diritti umani", 2021, p. 1193.

<sup>6</sup>SOMALVICO, M., *L'Intelligenza Artificiale*, Rusconi Editore, Milano, 1987, p.2.

si ispirano- seppur in maniera semplificata-al funzionamento del cervello umano e cercano di imitare i meccanismi dell'apprendimento.

Negli ultimi anni, l'intelligenza artificiale ha compiuto un ulteriore passo in avanti: non si limita più solo a “pensare”, ma ha acquisito la capacità di agire in modo autonomo all'interno di ambienti reali. Ne è un esempio significativo l'impiego di robot intelligenti, capaci di interagire con l'ambiente, adattarsi, prendere decisioni e, in certi contesti, comportarsi quasi come **entità indipendenti**, dotate di un certo grado di autosufficienza.<sup>7</sup>

## 1.2 L'IA come paradigma culturale secondo Grassi

Secondo Edmondo Grassi, l'intelligenza artificiale, nella sua forma attuale, non è solo un insieme di tecnologie computazionali, ma rappresenta un vero e proprio *paradigma culturale*, capace di trasformare il nostro modo di pensare, agire e relazionarci. Essa si presenta come una “protesi invisibile del pensiero”, influenzando non solo il piano tecnico, ma anche su quello affettivo, simbolico e sociale. Oggi, l'AI non si limita più a supportare le azioni umane, ma le rielabora, modificando la nostra percezione dello spazio, del tempo e dell'identità.<sup>8</sup>

In questo contesto, l'intelligenza artificiale emerge come una tecnologia trasversale e trasformativa che incide su molti aspetti della nostra vita, influenzando i processi educativi, lavorativi, relazionali e politici. La sua influenza non è affatto neutra; al contrario, è cruciale e cambia profondamente i valori e le scelte collettive.

Comprendere questa duplice natura dell'intelligenza artificiale, sia tecnica che culturale, è essenziale per affrontare con consapevolezza le numerose sfide che ci presenta. L'AI incide profondamente sul nostro modo di pensare, decidere e vivere insieme, ed è proprio per questo che solleva questioni complesse: etiche, giuridiche e sociali, che non possiamo più permetterci di ignorare.

---

<sup>7</sup>SOMALVICO, M., op.cit., p.15.

<sup>8</sup> GRASSI, E., *Etica e intelligenza artificiale*, Roma, Armando Editore, 2021, p. 19.

## **2. La Quarta rivoluzione industriale: evoluzione storica e ambiti di applicazione**

### **2.1 Una rivoluzione tecnologica e antropologica**

L'intelligenza artificiale ha giocato un ruolo chiave nella nascita della "Quarta Rivoluzione Industriale", un concetto teorizzato da Klaus Schwab<sup>9</sup> per descrivere l'attuale fase di trasformazione tecnologica. In questo periodo, i confini tra mondo fisico, digitale e biologico si stanno facendo sempre più sfumati ed interconnessi. L'AI non è solo una tecnologia utile, ma rappresenta un vero e proprio modello che ha cambiato e continua a trasformare radicalmente il modo in cui le persone interagiscono con l'ambiente, le istituzioni e tra di loro.<sup>10</sup>

Il suo impiego e sfruttamento nella vita quotidiana è in costante espansione. Dall'utilizzo di assistenti vocali alla personalizzazione delle esperienze di navigazione online, fino alla gestione automatizzata delle case intelligenti (smart home), l'AI si è silenziosamente ma inesorabilmente inserita nelle nostre abitudini quotidiane, diventando parte integrante della nostra routine, agendo spesso in maniera invisibile ma determinante. Essa influenza le nostre scelte, anticipa le preferenze e ci condiziona in modi che possono sembrare sottili, ma che sono profondamente radicati. Inoltre, filtra l'accesso alle informazioni, giocando un ruolo chiave nella formazione della nostra soggettività contemporanea in modo implicito e strutturale.<sup>11</sup>

### **2.2 Lavoro e mutazione delle competenze**

L'uso dell'AI nel mondo del lavoro ha iniziato come un supporto per le persone, aiutandole con compiti ripetitivi o poco impegnativi dal punto di vista cognitivo. Tuttavia, con l'avvento dell'automazione predittiva e dei sistemi di apprendimento automatico, l'IA ha cominciato a sostituire anche ruoli altamente qualificati: dalla valutazione del rischio, alla selezione del personale, dalla pianificazione industriale alle diagnosi

---

<sup>9</sup> SCHWAB, K., *La quarta rivoluzione industriale*, Milano, Franco Angeli, 2017, p. 15.

<sup>10</sup> CIPRIANI, A., GRAMOLATI, A., MARI, G (a cura di)., *Il lavoro 4.0. La Quarta Rivoluzione Industriale e le trasformazioni delle attività lavorative*, Firenze, Firenze University Press, 2018, pp. 9-11.

<sup>11</sup> CIPRIANI, A., GRAMOLATI, A., MARI, G (a cura di)., *op. cit.*, pp. 12-13

mediche, fino alla gestione delle risorse finanziarie. Questo ha portato ad una vera e propria rivoluzione nel mercato del lavoro e delle competenze richieste.<sup>12</sup>

Secondo un report di McKinsey citato da Pietro Piro, entro il 2030, fino al 15% della forza lavoro globale potrebbe essere influenzato da processi di sostituzione o trasformazione legati all'adozione dell'intelligenza artificiale. Tuttavia, non stiamo assistendo a una semplice "estinzione" del lavoro, ma piuttosto ad una sua evoluzione: mentre alcune mansioni scompaiono, altre emergono ed altre si fondono. La relazione uomo-macchina sta cambiando, passando da una logica di competizione a una logica di collaborazione (augmenting automation), dove la sinergia tra abilità umane e capacità computazionali apre nuove strade per efficienza, precisione e adattabilità. Questa rivoluzione non è solo tecnica, ma tocca profondamente la nostra umanità. Il lavoro, che storicamente ha rappresentato un pilastro dell'identità individuale e della cittadinanza sociale, si trova oggi a dover affrontare una crisi nel suo ruolo di collante simbolico e distributivo.<sup>13</sup>

### **2.3 Ecosistemi dell'innovazione e transdisciplinarietà**

Lo sviluppo dell'intelligenza artificiale rappresenta un modo di innovare che è davvero unico e strutturalmente nuovo. Tra le sue caratteristiche si distingue innanzitutto la transdisciplinarietà, che promuove l'integrazione di conoscenze che prima erano separate, rendendo fondamentale il passaggio di informazioni tra la ricerca teorica e le applicazioni pratiche. In tale contesto, si affermano ecosistemi di innovazione aperti, nei quali la cooperazione tra aziende, centri di ricerca, startup e istituzioni scientifiche avviene in modo fluido e dinamico, superando i confini tradizionali dei settori IT.

Le aziende più consolidate nei settori strategici stanno progressivamente adottando piani di innovazione agili, capaci di unire visione a lungo termine e flessibilità per affrontare la rapida evoluzione tecnologica. Questo processo coinvolge una varietà di attori, provenienti non solo dall'informatica o dall'ingegneria, ma anche da una molteplicità di ambiti applicativi, discipline scientifiche e segmenti sociali. Ne consegue la necessità

---

<sup>12</sup> PIRO, P., *Intelligenza artificiale (IA) e mercati del lavoro*, Preprint, ResearchGate, maggio 2025, p .5.

<sup>13</sup>Ibidem

di profili professionali più ampi, in grado di integrare competenze tecniche, etiche, relazionali e culturali. Infine, l'innovazione dell'IA si sviluppa su più livelli istituzionali, andando oltre le tradizionali strutture nazionali e affermandosi in un sistema multilivello, che coinvolge attori subnazionali e internazionali, in un contesto sempre più interconnesso e globale.<sup>14</sup>

## **2.4 AI nella sanità: la rivoluzione silenziosa**

Nel contesto sanitario, l'AI sta contribuendo in modo significativo a progressi come la medicina personalizzata, la diagnostica predittiva e nella creazione di piani terapeutici su misura. Come sottolinea il report *AI e Sanità* a cura della Fondazione Bassetti, l'adozione dell'intelligenza artificiale è un motore fondamentale del cambiamento nella trasformazione digitale della sanità odierna. Le innovazioni non stanno solo potenziando gli strumenti a disposizione dei professionisti, ma stanno anche modificando abitudini quotidiane, introducendo nuovi comportamenti prima sconosciuti, come l'uso di messaggistica istantanea, social network o il monitoraggio dei parametri biologici e trasformando profondamente pratiche consolidate, come quelle legate al lavoro, all'istruzione e alle professioni.

La medicina è oggi uno dei settori più coinvolti da questa spinta evolutiva. Secondo il rapporto *Artificial Intelligence and Life in 2030* dell'Università di Stanford, il settore sanitario è tra quelli in cui l'intelligenza artificiale avrà un impatto maggiore. In Italia, persistono però ostacoli infrastrutturali e normativi: la disomogeneità nella diffusione del fascicolo sanitario elettronico, le difficoltà di interoperabilità tra dati clinici provenienti da fonti diverse ed i vincoli legati alla normativa sulla privacy limitano l'attuazione di sistemi integrati e predittivi. Anche l'integrazione con dati generati dai pazienti stessi tramite dispositivi come cardiofrequenzimetri o app mediche appare ancora lontana, nonostante gli investimenti previsti dal PNRR. Tuttavia, applicazioni avanzate come la robotica chirurgica e la stampa 3D sono già presenti in molte strutture ospedaliere. Da qui

---

<sup>14</sup>PIRO, P., op.cit., p.6.

l'urgenza di incentivare un diffusa alfabetizzazione digitale, in grado di accompagnare con consapevolezza e spirito critico l'adozione di queste tecnologie.<sup>15</sup>

In futuro, l'intelligenza artificiale potrà potenziare ulteriormente l'efficacia del sistema sanitario, grazie alla capacità di analizzare grandi volumi di dati, aumentare la precisione diagnostica e personalizzare i percorsi di cura. Algoritmi sofisticati sono in grado di individuare precocemente l'insorgenza di malattie, suggerendo trattamenti mirati. Tecnologie immersive come la realtà virtuale (VR) e aumentata (AR) sono già impiegate nella formazione dei medici e nella pianificazione degli interventi chirurgici, offrendo esperienze simulate sempre più realistiche. Dispositivi indossabili e sensori biometrici monitorano in tempo reale i parametri vitali, promuovendo un approccio alla salute predittivo e continuo. Le nanotecnologie permettono il rilascio mirato di farmaci attraverso nanoparticelle intelligenti, mentre la stampa 3D rende possibile la creazione di protesi su misura, modelli anatomici e tessuti biologici sperimentali. La telemedicina, infine, riduce le disuguaglianze territoriali nell'accesso alle cure, mentre la digitalizzazione delle cartelle cliniche garantisce una gestione sicura e integrata della storia medica dei pazienti, favorendo diagnosi più rapide e trattamenti più efficaci.<sup>16</sup>

## 2.5 Difesa e sicurezza strategica

Nel settore della difesa<sup>17</sup>, l'intelligenza artificiale sta rivoluzionando profondamente le dinamiche operative e strategiche. L'impiego di sistemi autonomi, come i droni intelligenti utilizzati per missioni di sorveglianza o attacco, testimonia una trasformazione già in atto nella gestione dei conflitti. Accanto a questi, tecnologie all'avanguardia come le armi ipersoniche e a energia diretta stanno ridefinendo il volto stesso della guerra.

Perfino l'organizzazione logistica militare beneficia dell'adozione dell'intelligenza artificiale, l'automazione dei processi di approvvigionamento sul campo consente una gestione più efficiente delle risorse in contesti operativi complessi. Parallelamente,

---

<sup>15</sup> FONDAZIONE BASSETTI, *AI e Sanità. Tecnologie, etica, responsabilità*, in LIMONE, D. A. (a cura di), *AI & Human Rights*, Roma, 2023, pp. 53-54.

<sup>16</sup>PIRO, P., op.cit., pp.7-8.

<sup>17</sup> Tutti i contenuti della presente sezione, dal settore della difesa fino alla regolazione economica inclusa, sono tratti, salvo diversa indicazione da PIRO, P., *Intelligenza artificiale (IA) e mercati del lavoro*. Preprint, ResearchGate, maggio 2025: pp.11-16.

l'utilizzo della realtà virtuale e aumentata nei programmi di addestramento garantisce l'esperienza di simulazione estremamente realistiche.

Numerose di queste innovazioni, nate per scopi puramente bellici, presentano una natura duale e trovano applicazione anche nel contesto civile, alimentando un dibattito etico e geopolitico sempre più urgente sulla convergenza tra sfera militare e tecnologica.

## **2.6 Ingegneria e sostenibilità ambientale**

Nel campo dell'ingegneria, l'impatto trasformativo dell'intelligenza artificiale è sempre più evidente. Si assiste a un proliferare di soluzioni che uniscono progresso tecnologico e rispetto per l'ambiente, come le reti intelligenti per distribuire l'energia in modo smart e le infrastrutture sostenibili, ad esempio l'agricoltura verticale. Progetti ambiziosi come l'hyperloop o i droni per l'impollinazione mostrano come l'ingegneria moderna non si occupi solo di costruire, ma anche di farsi carico di delineare scenari futuri inediti in questo contesto, la sinergia tra intelligenza artificiale, ispirazione biologica e progettazione sistematica diventa essenziale.

## **2.7 Agricoltura intelligente e tecnologie sostenibili**

Pure il mondo dell'agricoltura ha visto l'arrivo dell'intelligenza artificiale, aprendo la strada ad una conduzione di campi sempre più accurata, rispettosa dell'ambiente ed automatizzata. Grazie a droni e sensori, si tiene sotto controllo costantemente la salute delle piante e speciali programmi informatici migliorano l'utilizzo di risorse preziose come l'acqua ed i fertilizzanti. L'agricoltore di oggi diventa quasi un esperto potenziato, non è soltanto capace di captare e gestire informazioni complesse, ma riesce a utilizzare e controllare sistemi di produzione tecnologici ed intelligenti.

## **2.8 Logistica e automazione della filiera produttiva**

Nel campo della logistica, l'intelligenza artificiale rappresenta una leva decisiva per l'automazione e l'ottimizzazione della catena del valore. L'introduzione di veicoli autonomi, robot mobili e sensori IoT consente il tracciamento continuo delle merci, riducendo i tempi di transito e migliorando la sicurezza. Le piattaforme collaborative integrano dati in tempo reale, facilitando decisioni tempestive lungo l'intera filiera.

Queste trasformazioni ridisegnano l'organizzazione del lavoro, potenziano l'efficienza operativa e pongono nuove sfide in termini di governance, standard e interoperabilità dei sistemi.

## **2.9 Finanza, banche e moneta digitale**

Il settore bancario vive una fase di riconfigurazione profonda, guidata dall'adozione di strumenti intelligenti capaci di personalizzare i servizi e rispondere alle esigenze specifiche di ogni singolo cliente, aumentando esponenzialmente la sicurezza dei sistemi e semplificando al contempo le operazioni finanziarie. Si ricorre ad algoritmi predittivi per la gestione del rischio, la concessione del credito e la prevenzione delle frodi. La tecnologia blockchain introduce nuovi modelli di tracciabilità e fiducia. Contemporaneamente, il progetto dell'euro digitale solleva interrogativi rilevanti in merito all'autonomia strategica, alla privacy finanziaria e al ruolo futuro delle banche nel contesto della moneta pubblica digitale.

## **2.10 Istruzione e ambienti di apprendimento digitali**

Nel mondo dell'istruzione, l'intelligenza artificiale si configura come agente di cambiamento profondo. L'espansione di contesti didattici digitali, insegnanti virtuali e piattaforme personalizzate modifica il modo in cui ci avviciniamo al sapere, superando le barriere spazio-temporali. La realtà aumentata e quella virtuale offrono esperienze immersive, mentre l'analisi dei dati scolastici apre la strada a percorsi di formazione creati su misura. Malgrado ciò, questo processo richiede una riflessione critica sul rapporto tra innovazione e qualità dell'apprendimento, evitando una deriva tecnocratica che sacrifichi il rapporto umano nell'educazione.

## **2.11 Aerospazio e competizione strategica globale**

Nel settore aerospaziale, l'IA sta ridefinendo tanto la progettazione quanto la gestione operativa. Si usano sistemi basati su modelli generativi per migliorare al massimo l'aerodinamica e tagliare i consumi, ed allo stesso tempo la manutenzione predittiva degli aeromobili consente un controllo permanente e anticipatorio rispetto alle eventuali criticità. L'AI trova applicazione anche nella gestione del traffico aereo e per organizzare missioni satellitari davvero complicate. In Italia, l'adozione di strategie finanziarie

innovative e l'espansione dei poli industriali dimostrano che si vuole rafforzare la competitività nazionale in un settore sempre più strategico e mondiale.

## **2.12 Governance pubblica e regolazione algoritmica**

L'intelligenza artificiale sta assumendo un ruolo crescente anche nei meccanismi di regolazione e governance economica. Sistemi automatizzati vengono utilizzati per identificare frodi, ottimizzare la raccolta fiscale e monitorare gli indicatori di rischio sistemico. Tuttavia, l'adozione di strumenti decisionali algoritmici nel settore pubblico solleva interrogativi cruciali in termini di trasparenza, controllo democratico e rispetto dei diritti fondamentali. La sfida sarà coniugare efficienza e giustizia, innovazione e responsabilità.

## **3. Il dibattito etico**

### **3.1 l'immaginario collettivo tra utopia e distopia**

l'IA è spesso fraintesa, anche a causa di narrazioni fantascientifiche che, pur avendo anticipato molte sue potenzialità, ne hanno alimentato una visione distorta o eccessivamente antropomorfizzata.

Nel suo testo "Etica e Intelligenza Artificiale", Grassi sottolinea come l'AI venga spesso proiettata in una dimensione immaginaria, a metà tra la promessa e la minaccia, in questo duello costante tra il bene ed il male. Questo dualismo richiama il mito di Prometeo, figura archetipica del dominio umano sulla natura tramite la tecnica, e rivela quanto la nostra visione dell'intelligenza artificiale sia ancora intrisa di simbolismi culturali e paure ancestrali.<sup>18</sup>

In realtà, l'intelligenza artificiale è una disciplina complessa, che solleva interrogativi **etici, sociali e giuridici** di crescente attualità.

---

<sup>18</sup> GRASSI, E., *Etica e intelligenza artificiale. Questioni aperte*, Tesi di Dottorato in Ricerca Sociale Teorica e Applicata, Università degli Studi Roma Tre, XXXII ciclo, tutor Prof.ssa Marina D'Amato, 2020, p. 19.

### 3.2 Coscienza artificiale e dignità umana: verso un'etica della regolazione

Uno dei tratti più affascinanti e sorprendenti dell'IA è la sua straordinaria capacità di evolversi incessantemente. Ciò che oggi viene considerato come intelligente, innovativo o addirittura rivoluzionario, domani potrebbe risultare obsoleto. L'IA, dunque, si configura come una disciplina di confine, in perenne ridefinizione e movimento, che accompagna e talvolta anticipa il progresso scientifico e tecnologico.

Questa evoluzione continua solleva, e continuerà a sollevare, interrogativi che travalicano l'ambito tecnico per toccare in profondità la sfera filosofica. Se si accetta l'assunto razionalista cartesiano secondo cui *cogito, ergo sum*<sup>19</sup> – ovvero, che l'autocoscienza scaturisce dalla capacità di pensare – allora si potrebbe ipotizzare, almeno teoricamente, che sistemi artificiali dotati di enorme capacità computazionale siano destinati, prima o poi, a sviluppare una forma di autocoscienza. Tuttavia, sia Edmondo Grassi che Laura Palazzani mettono in guardia da questa equivalenza semplicistica. Il primo sottolinea che la coscienza non è semplicemente un prodotto del calcolo, bensì un'esperienza incarnata, emotiva e relazionale.<sup>20</sup> La seconda osserva come l'identità umana si fonda su elementi irriducibili all'informazione: dignità, libertà, interiorità. Se la macchina può imitare, non per questo può *esperire*. Può replicare il comportamento ma non viverlo.<sup>21</sup>

Secondo Palazzani, l'intelligenza artificiale appartiene al mondo del “fare” (*doing*), mentre l'essere umano appartiene al mondo dell'“essere” (*being*): la macchina funziona; l'uomo esiste. L'essere umano è consapevole di sé, dotato di interiorità, libero arbitrio e capace di dare significato alla propria esperienza. La macchina, per quanto sofisticata, rimane priva di un'identità unitaria e di coscienza morale.<sup>22</sup>

---

<sup>19</sup> DESCARTES, R., *Discorso sul metodo*, parte IV, 1637; ed. it. a cura di E. Garin, Laterza, Bari, 2009.

<sup>20</sup> GRASSI, E., op. cit., p. 23.

<sup>21</sup> PALAZZANI, L., *Etica della regolazione dell'AI*, in *Etica per le tecnologie emergenti*, CNR Edizioni, Roma, 2023, p. 16.

<sup>22</sup> Ivi, p.18.

Palazzani sottolinea, inoltre, come l'intelligenza artificiale, nella sua forma attuale, non sia una tecnologia neutra: essa incarna una precisa visione antropologica e solleva dilemmi etici e giuridici complessi. Il cuore del problema sembra risiedere nella tensione tra due approcci contrapposti: da un lato, quello tecno-centrico, che promuove il superamento dei limiti umani attraverso l'artificializzazione dell'umano e l'espansione post-umana;<sup>23</sup> dall'altro, l'approccio umano-centrico, che ribadisce l'unicità e l'irriducibilità dell'essere umano, in quanto dotato di autocoscienza, libertà, affettività e relazionalità.<sup>24</sup>

Per evitare derive disumanizzanti, Palazzani propone un'etica normativa che guidi lo sviluppo dell'IA lungo direttrici chiare: *controllo umano significativo, affidabilità dei sistemi, trasparenza e spiegabilità delle logiche algoritmiche, inclusività dei dati e tracciabilità delle decisioni*.<sup>25</sup> Tali requisiti sono considerati indispensabili per evitare discriminazioni, lesioni alla privacy e fenomeni di deresponsabilizzazione morale, assicurando al contempo il rispetto della dignità umana nel contesto digitale.

A partire da questa tensione si pongono le principali questioni etiche e giuridiche connesse all'impiego dell'IA: il rischio di discriminazioni algoritmiche, la perdita di privacy, la difficoltà di attribuire responsabilità giuridica a decisioni autonome, la manipolazione dell'opinione pubblica, il controllo sociale e, più in generale, il pericolo di una progressiva reificazione della persona. Come osserva Palazzani, non si tratta solo di governare le applicazioni dell'IA, ma di interrogarsi sul modello antropologico che ne guida lo sviluppo.<sup>26</sup>

Sul piano giuridico, emerge con forza la necessità urgente di norme vincolanti che tutelino i diritti fondamentali da automatismi ciechi. Tuttavia, la regolamentazione, come evidenzia la stessa Palazzani, deve anche promuovere le potenzialità positive

---

<sup>23</sup> Ivi, pp. 18-19.

<sup>24</sup> Ibidem

<sup>25</sup> Ivi, p.20.

<sup>26</sup> PALAZZANI, L., op. cit., p. 22.

dell'innovazione, adottando un approccio equilibrato e prudentiale, capace di valutare benefici e rischi e di adattarsi all'evoluzione tecnologica.<sup>27</sup>

Solo attraverso un equilibrio maturo tra etica e diritto, tra coscienza morale e norma giuridica potremmo immaginare un'AI realmente al diritto dell'uomo. Non si tratta di arrestare il progresso, ma di orientarlo nella direzione giusta: quella che riconosce e protegge il valore umano.

#### **4. L'algoetica: una nuova etica per i sistemi intelligenti**

L'evoluzione dell'intelligenza artificiale ha inciso profondamente sul rapporto tra esseri umani e tecnologia, generando questioni che vanno ben oltre le questioni tecniche o giuridiche sollevando interrogativi che vanno ben oltre le mere questioni tecniche o giuridiche. Come già osservato nella parte precedente, le sfide poste dall'AI non si esauriscono nella responsabilità per atti compiuti da sistemi autonomi, né si limitano alla necessità di assicurarne trasparenza e non discriminazione. Piuttosto, richiedono, una profonda riflessione etica radicale sul nostro modo di essere, sulle nostre relazioni e sulle dinamiche di potere che attraversano la nuova era della società digitale.

È in questo contesto che si colloca il nuovo concetto di **algoetica**, ovvero una disciplina etica autonoma che studia i principi, i limiti e i valori che dovrebbero guidare la progettazione, l'uso e la regolazione degli algoritmi intelligenti.<sup>28</sup>

Il termine “algoetica” è stato proposto da Luciano Floridi all'interno del progetto europeo AI4People, con l'intento di sottolineare la necessità di un nuovo approccio etico pubblico che orienti lo sviluppo degli algoritmi a partire da valori condivisi come la dignità, la giustizia, l'autonomia, la beneficenza e la responsabilità.<sup>29</sup>

Secondo Floridi, l'intelligenza artificiale rappresenta una nuova forma di “*smart agency*”, ovvero un agente intelligente capace di influenzare attivamente gli ambienti sociali,

---

<sup>27</sup> Ivi, p. 24.

<sup>28</sup> FLORIDI, L., et al., *AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*, in *Minds and Machines*, vol. 28, 2018, p. 689 ss.

<sup>29</sup> FLORIDI, L., et al., op. cit., p. 690.

relazionali e ambientali, rimodellando i confini dell'agire umano: “*AI is a powerful force, a new form of smart agency, which is already reshaping our lives, our interactions, and our environments*”.<sup>30</sup>

Proprio per questo, l'AI non può essere trattata come una tecnologia neutra, né può essere regolata esclusivamente a posteriori mediante strumenti normativi reattivi. In questa prospettiva, l'etica algoritmica – o algoretica – non è solo un codice esterno alla tecnica, ma come un'etica “*nella tecnica*”, integrata nei processi di calcolo fin dalla fase di progettazione.<sup>31</sup>

L'obiettivo dichiarato di AI4People è quello di passare “*from AI ethics to Ethics by design*”, ovvero incorporare i valori fondamentali direttamente nelle architetture decisionali e nei sistemi intelligenti,<sup>32</sup> rendendo l'etica parte costitutiva dell'algoritmo stesso.

Il documento individua cinque principi chiave: beneficenza, non maleficenza, autonomia, giustizia ed esplicabilità, da applicare in modo sistemico e concreto lungo durante l'intero ciclo di vita dell'intelligenza artificiale.<sup>33</sup> L'algoritmo, in questo senso, rappresenta non solo una risposta teorica, ma anche una strategia operativa per affrontare le trasformazioni che l'AI impone alla società contemporanea, promuovendo un'azione regolativa anticipatoria, trasparente e centrata sull'umano.<sup>34</sup>

Sulla base di queste premesse, nel prosieguo del capitolo esamineremo, in primo luogo, le principali problematiche sollevate dall'impiego dell'intelligenza artificiale nei contesti sociali, economici e giuridici. In secondo luogo, analizzeremo come ciascuna di queste criticità possa essere affrontata alla luce dei valori fondamentali proposti dall'algoritmo, con l'obiettivo di valutare la reale efficacia del paradigma etico delineato da AI4People.

---

<sup>30</sup> Ibidem.

<sup>31</sup> Ivi, p. 692.

<sup>32</sup> Ivi, p. 696.

<sup>33</sup> Ivi, p. 697.

<sup>34</sup> Ivi, p. 701.

## 4.1 Spiegabilità e dignità umana

Nel dibattito contemporaneo sull'intelligenza artificiale, due temi si impongono con particolare urgenza e rilevanza: la spiegabilità dei processi decisionali automatizzati e la tutela della dignità umana. Si tratta di questioni profondamente connesse. Comprendere come e perché un sistema di IA giunga a una determinata decisione rappresenta, infatti, il primo passo per garantire il rispetto dell'individuo nella sua autonomia morale e nella sua unicità.

Come osserva Giampaolo Collecchia nel suo contributo sull'IA spiegabile, molti dei sistemi oggi più avanzati – in particolare quelli basati su **reti neurali profonde** – operano secondo logiche interne così complesse da risultare opache, dando origine al cosiddetto modello a scatola nera (*black box model*), in cui è possibile osservare l'input e l'output, ma non il processo intermedio che conduce alla decisione finale. Questa opacità non rappresenta solo un limite tecnico: essa genera un vuoto in termini di responsabilità, affidabilità e controllo umano, minando la fiducia nella tecnologia, soprattutto in settori estremamente delicati come la sanità, la giustizia, la finanza, la mobilità e il lavoro <sup>35</sup>

La disciplina emergente della *eXplainable Artificial Intelligence* (XAI) nasce proprio per rispondere a questa sfida, con l'obiettivo di rendere interpretabili e comprensibili i meccanismi interni degli algoritmi. <sup>36</sup>Tuttavia, la questione non si esaurisce in un problema tecnico: rendere spiegabili i sistemi automatizzati significa riconoscere e garantire un diritto fondamentale alla comprensione, condizione necessaria per l'autodeterminazione dell'individuo. Questo principio è stato accolto anche dalle istituzioni europee, che nei documenti strategici programmatici dal Libro Bianco della Commissione al Regolamento (UE) 2024/... (AI Act), hanno identificato la trasparenza e la spiegabilità come requisiti imprescindibili per uno sviluppo dell'AI compatibile con i diritti umani. L'art. 50 dell'AI Act impone obblighi espliciti ai fornitori e agli utilizzatori

---

<sup>35</sup> COLLECCHIA, G., *Apriamo la scatola nera: l'intelligenza artificiale spiegabile*, in *Recenti Progressi in Medicina*, 2021, n. 11.

<sup>36</sup> Cfr. AI4PEOPLE, *An Ethical Framework for a Good AI Society*, in *Minds and Machines*, vol. 28, n. 4, 2018, pp. 689-707. Si veda anche COMMISSIONE EUROPEA, *Linee guida etiche per un'IA affidabile*, Bruxelles, aprile 2019.

affinché sia chiaro, identificabile e verificabile il carattere artificiale dei contenuti generati, soprattutto in caso di *deep fake* o decisioni ad elevato impatto<sup>37</sup>

Tuttavia, comprendere l'opacità dei sistemi non significa semplicemente ispezionarne il codice sorgente. Come sottolinea Martoni, la spiegabilità non può ridursi ad un requisito tecnico: essa rappresenta un metodo per interrogare l'innovazione alla luce dell'etica. L'intelligenza artificiale non è un soltanto un mero strumento meccanicistico: è un "*fatto sociale totale*", capace di riplasmare interi ambienti di vita. Da qui, l'urgenza di stabilire, fin dalle fondamenta, i principi etici che devono orientarne la progettazione e l'utilizzo, come la trasparenza, la tracciabilità, la responsabilità e il rispetto della dignità umana<sup>38</sup>

In questo senso, il diritto alla spiegazione non può essere separato da ciò che rappresenta il suo fondamento più profondo: la dignità della persona. Quest'ultima costituisce il pilastro nell'intero impianto dei diritti fondamentali e costituisce un limite invalicabile oltre il quale nessuna innovazione può legittimamente spingersi. Questa opacità ha ripercussioni gravi soprattutto quando l'IA viene impiegata in ambito giurisdizionale. La mancanza di una motivazione comprensibile – sancita dall'art. 111 della Costituzione – rende la decisione automatizzata difficilmente contestabile e conferisce loro una parvenza di oggettività e neutralità spesso infondata.

Per queste ragioni, la spiegabilità deve essere considerata un diritto autonomo, distinto dal semplice accesso ai dati, e deve essere garantita anche nei casi in cui l'intervento umano sia solo apparente. Il principio di non esclusività algoritmica può dirsi rispettato solo quando il decisore umano è in grado di fornire una motivazione consapevole e comprensibile dell'adesione al risultato generato dal sistema automatizzato<sup>39</sup>

---

<sup>37</sup> Regolamento (UE) 2024/... del Parlamento europeo e del Consiglio che stabilisce regole armonizzate in materia di intelligenza artificiale (AI Act), art. 50. Sul punto, cfr. anche Commissione Europea, *Proposta di regolamento sull'intelligenza artificiale*, COM(2021) 206 final.

<sup>38</sup> MARTONI, M., *Spiegare le pieghe. Alcune osservazioni su trasparenza e spiegabilità*, in *L'Ircocervo*, n. 1, 2024, pp. 215–217.

<sup>39</sup> MACERATINI, A., *Dati digitali e intelligenza artificiale. Sfide etiche e giuridiche*, in *Tigor. Rivista di scienze della comunicazione e di argomentazione giuridica*, a. XIV, n. 2, 2022, p. 69.

Quando l'intelligenza artificiale diventa strumento di esclusione, sorveglianza, standardizzazione o spersonalizzazione, entrambe queste dimensioni, spiegabilità e dignità, vengono messe profondamente in discussione. Come già evidenziato da Collecchia, il rischio è che l'individuo venga ridotto a un profilo statistico o a un oggetto computazionale, privato della possibilità di comprendere o contestare ciò che lo riguarda. Difendere la dignità in questo contesto significa garantire libertà decisionale piena, consapevole e informata, riaffermando la centralità dell'essere umano come fine, e mai come mezzo.

In tale contesto, non sorprende che uno dei primi strumenti giuridici ad occuparsi esplicitamente dei processi decisionali automatizzati sia stato proprio il Regolamento (UE) 2016/679 (GDPR). L'articolo 22 di tale Regolamento riconosce il diritto dell'interessato a non essere sottoposto a una decisione basata unicamente su un trattamento automatizzato, inclusa la profilazione, qualora tale decisione produca effetti giuridici rilevanti o incida in modo significativo sulla persona.

Come osservato anche da Maceratini, il GDPR pone così le basi per una spiegabilità giuridicamente rilevante, sancendo il diritto dell'individuo a ricevere una spiegazione comprensibile delle logiche sottese al processo decisionale. In questo senso, il Regolamento anticipa alcuni dei principi che oggi trovano piena formalizzazione nel nuovo AI Act, delineando un'architettura normativa fondata sull'intreccio tra dignità, trasparenza e la responsabilità nei sistemi automatizzati.<sup>40</sup>

## **4.2 Non discriminazione algoritmica**

L'intelligenza artificiale, pur nella sua potenziale capacità di rendere le decisioni più rapide ed efficienti, non è affatto immune da possibili distorsioni sistemiche. Contrariamente all'idea ancora diffusa di una tecnologia neutrale, molteplici studi dimostrano che gli algoritmi sono profondamente influenzati dalla tipologia di dati su cui vengono addestrati e dalle scelte progettuali che li definiscono. Il rischio principale, come sottolineato da autorevoli studiosi, non risiede solo nel fatto che l'AI "*apprende*" da un

---

<sup>40</sup> Ivi, p. 67.

passato già segnato da disuguaglianze ed ingiustizie sociali, ma anche nella sua tendenza a riprodurlo ed istituzionalizzarlo come se fosse un modello imparziale<sup>41</sup>.

Uno dei meccanismi più subdoli con cui si manifesta la discriminazione algoritmica è l'impiego di dati apparentemente neutri – i cosiddetti *proxy* – che risultano però fortemente correlati con caratteristiche protette come l'origine etnica, il genere o la condizione socio-economica. Anche se tali dati, non si riferiscono direttamente a categorie sensibili, possono replicarne gli effetti in modo sistematico, generando discriminazioni che sfuggono ai controlli tradizionali.<sup>42</sup> Il risultato è che soggetti già svantaggiati finiscono per essere penalizzati da sistemi che si dichiarano formalmente equi, ma che, nei fatti perpetuano esclusione e disuguaglianza.<sup>43</sup>

Un caso emblematico di questa dinamica è rappresentato dai sistemi algoritmici impiegati nelle piattaforme di food delivery. I rider vengono valutati da software che analizzano parametri come puntualità, tasso di accettazione delle consegne e disponibilità oraria. Sebbene questi criteri appaiano neutri ed oggettivi, generano in realtà una segmentazione invisibile ma profonda, che finisce per escludere progressivamente quei lavoratori che, per ragioni legate alla residenza, alla condizione familiare o ad altri vincoli soggettivi, non possono garantire performance “ottimali”.<sup>44</sup> Si crea una vera e propria gerarchia algoritmica, in cui chi parte da condizioni svantaggiate viene ulteriormente penalizzato, anche economicamente, fino a rischiare, nei casi più gravi, l'esclusione automatizzata dalla piattaforma<sup>45</sup>.

IA ciò si aggiunge un ulteriore elemento critico: il lavoratore, non è messo in condizione di comprendere il funzionamento dell'algoritmo, né tantomeno di contestarne le decisioni. Ciò è aggravato dal fatto che molte piattaforme considerano il software come

---

<sup>41</sup>TOMMASI, S., *Algoritmi e discriminazione: la responsabilità dell'umano nella macchina*, in ALTROCONSUMO (a cura di), *AI e Discriminazione*, Altroconsumo Edizioni, Milano, 2020, pp. 13–14

<sup>42</sup> TOMMASI, S., op. cit., pp. 13–14.

<sup>43</sup> Ibidem.

<sup>44</sup> Ivi, pp.21–23

<sup>45</sup> TOMMASI, S., op. cit., p. 23.

un sistema chiuso e proprietario, sottraendolo al controllo pubblico. Si realizza così una governance algoritmica opaca, il cui il potere decisionale sfugge sia al lavoratore che al legislatore<sup>46</sup>.

Il **Regolamento Generale sulla Protezione dei Dati (GDPR)** ha cercato di porre un argine a queste dinamiche. In particolare, l'articolo 22 riconosce il diritto di ogni individuo a non essere sottoposto a decisioni basate unicamente su trattamenti automatizzati. Tuttavia, come osservato dalla dottrina, questa tutela si rivela spesso inadeguata per i casi di discriminazione algoritmica indiretta. Le eccezioni all'articolo 22 sono numerose, la definizione di decisione "unicamente automatizzata" è restrittiva e le modalità di esercizio dei diritti restano incerte.

Il regolamento Generale sulla protezione dei dati personali (GDPR), come abbiamo già evidenziato, ha tentato di contrastare i rischi legati all'automazione decisionale, riconoscendo il diritto all'individuo a non essere sottoposto a decisioni fondate unicamente su trattamenti automatizzati.<sup>47</sup> Tuttavia, anche le altre disposizioni fondamentali del Regolamento incontrano limiti significativi.

L'articolo 5, che stabilisce i principi di liceità, correttezza e trasparenza del trattamento, si scontra con l'opacità strutturale dei sistemi algoritmici proprietari, spesso sottratti ad un reale scrutinio pubblico.<sup>48</sup>

L'articolo 9, che vieta il trattamento dei dati sensibili salvo eccezioni, non riesce a prevenire in modo diretto efficace l'uso indiretto di proxy discriminatori. Anche in questo caso, come già visto a proposito della spiegabilità algoritmica, il GDPR si dimostra uno strumento giuridico ancora troppo fragile per affrontare con efficacia le nuove sfide poste dall'intelligenza artificiale.<sup>49</sup>

In definitiva, l'attuale impianto normativo appare insufficiente a contrastare fenomeni strutturali di esclusione algoritmica. Come dimostrato dal caso delivery, la

---

<sup>46</sup> Ivi, p.15.

<sup>47</sup> Art. 22, Regolamento (UE) 2016/679 (GDPR)

<sup>48</sup> Ivi, art. 5.

<sup>49</sup> Art. 9, Regolamento (UE) 2016/679 (GDPR).

discriminazione non è il frutto di un errore accidentale, ma l'esito sistemico di modelli di ottimizzazione che ignorano la giustizia sostanziale. Occorrono strumenti capaci di intercettare la discriminazione indiretta, imporre audit preventivi, e garantire trasparenza ex ante sulle logiche decisionali, restituendo centralità alla persona anche nell'era dell'intelligenza artificiale.

### **4.3 Il controllo umano significativo**

Quando l'intelligenza artificiale entra in gioco nei processi decisionali ad alto impatto, diventa necessario ripensare profondamente ad i concetti di responsabilità, legittimità e giustizia procedurale. In questo contesto, il principio del controllo umano significativo si configura come un pilastro essenziale: una garanzia che l'autonomia tecnica dei sistemi algoritmici non comprometta la centralità dell'essere umano. Il coinvolgimento umano, infatti, non può limitarsi ad un ruolo simbolico o ad una vigilanza formale dunque supervisione passiva. Deve essere concreto, informato e attivo in grado di incidere concretamente sulle scelte automatizzate in ogni fase del ciclo di vita dell'AI<sup>50</sup>

Perché sia davvero significativo, il controllo umano deve assicurarsi che l'essere umano resti *meaningfully in the loop*, cioè inserito nel processo decisionale con consapevolezza, competenza e potere di intervento. Non basta solo assistere al funzionamento dell'algoritmo: bisogna essere in grado di sospendere un processo ritenuto ingiusto, correggerne la direzione, comprendere le logiche e risalire chiaramente alla catena delle responsabilità. L'assenza di questi presidi produce un pericoloso vuoto normativo ed etico, in cui le scelte compiute da macchine vengono accettate acriticamente, come se fossero inevitabili o neutre <sup>51</sup>.

Il tema diventa ancora più urgente quando l'intelligenza artificiale viene utilizzata in ambiti che incidono sui diritti fondamentali delle persone: la giustizia, il lavoro, i servizi pubblici, la sanità. In questi settori, l'esclusione del fattore umano dal processo decisionale non è ammissibile. Il rischio è di costruire sistemi opachi, inaccessibili, e spesso discriminatori. A questo si aggiunge un ulteriore pericolo: anche quando il

---

<sup>50</sup> UNESCO – COMEST, *Ethical Principles for the Use of Artificial Intelligence*, Paris, UNESCO, 2021, p. 9.

<sup>51</sup> UNESCO – COMEST, op. cit., p. 10.

controllo umano è formalmente previsto, esso potrebbe essere affidato a soggetti non adeguatamente formati, incapaci di comprendere e valutare criticamente delle competenze necessarie per interpretare e valutare le decisioni dell’algoritmo. In tal caso, la supervisione diventerebbe un atto privo di efficacia.<sup>52</sup>

Perché si possa parlare di controlli umano effettivo occorrono precise *condizioni abilitanti*: la trasparenza delle logiche decisionali, la tracciabilità dei dati, l’accessibilità ai criteri impiegati, la possibilità di ricorso e di rettifica. Senza questi strumenti, ogni forma di controllo resta illusoria, e l’efficienza tecnologica rischia di oscurare la giustizia sostanziale<sup>53</sup>.

Nelle riflessioni internazionali più avanzate, il controllo umano è inteso non solo nella sua dimensione tecnica: ma come garanzia democrazia, tra automazione e Stato di diritto. Non si tratta di frenare il progresso, ma di orientarlo secondo una visione antropocentrica, in cui l’essere umano mantiene il potere di decidere sulla base di valori morali, contesto sociale, empatia e senso di equità. In questa prospettiva, l’intelligenza artificiale non è un soggetto autonomo, ma uno strumento da regolamentare, cui applicare principi giuridici e responsabilità condivise<sup>54</sup>.

Difendere il principio del controllo umano significa, in ultima analisi, riaffermare che anche nell’epoca dell’automazione, la sovranità decisionale appartiene all’uomo. Solo a partire da questa consapevolezza sarà possibile costruire una governance algoritmica realmente inclusiva, che sappia coniugare efficienza e giustizia, innovazione e trasparenza, precisione tecnica e valore umano.<sup>55</sup>

#### **4.4 La responsabilità nell’era dell’IA**

Tra le questioni più complesse sollevate dall’intelligenza artificiale vi è quella, tutt’altro che risolta, della responsabilità. L’IA, infatti, agisce sulla base di processi di apprendimento che possono evolvere autonomamente e sfuggire al controllo diretto

---

<sup>52</sup> Ivi, p. 11.

<sup>53</sup> Ivi, p. 12.

<sup>54</sup> Ivi, p. 14.

<sup>55</sup> UNESCO – COMEST, op. cit., p. 15.

dell'essere umano. Questo rende difficile, se non impossibile applicare i tradizionali schemi di imputazione della responsabilità propri del diritto civile.

La crescente diffusione di sistemi di intelligenza artificiale dotati di autonomia operativa e capacità adattiva solleva interrogativi radicali: chi risponde di una decisione algoritmica non prevista? Chi ne sopporta le conseguenze giuridiche, morali ed economiche?

Gli attuali modelli di imputazione fondati su colpa o dolo risultano spesso inadeguati a fronte di comportamenti algoritmici non deterministici, generati da sistemi complessi che operano in ambienti dinamici e interconnessi. In questo contesto, si impone un ripensamento profondo delle categorie giuridiche classiche, con l'obiettivo di bilanciare l'innovazione tecnologica con la tutela effettiva dei diritti dei soggetti coinvolti, specialmente quando si verifica un danno generato da sistemi complessi o autonomi.

Ugo Ruffolo, in un contributo particolarmente incisivo, evidenzia come le architetture basate su machine learning e deep learning abbiano totalmente stravolto il nesso causale tra condotta, colpa e danno, rendendo necessario un cambio di paradigma giuridico<sup>56</sup>. Secondo l'autore, la responsabilità non può più concentrarsi sull'agente umano, ma deve essere redistribuita lungo tutta la filiera di sviluppo ed utilizzo del sistema algoritmico, valorizzando il principio di accountability in chiave multilivello.<sup>57</sup>

L'obiettivo non è solo quello di attribuire un risarcimento ex post, ma di imporre doveri ex ante di trasparenza, controllo, documentazione e tracciabilità, ispirandosi al modello del GDPR e ai contenuti della proposta europea dell'AI Liability Directive<sup>58</sup>. Ruffolo sottolinea l'importanza della "responsabilità funzionale" e dell'adattamento del diritto civile codificato, senza cedere alla tentazione di una normazione ipertrofica e frammentaria<sup>59</sup>. Tra le soluzioni prospettate si evidenziano l'adozione di sistemi

---

<sup>56</sup>RUFFOLO, U. (a cura di), *Intelligenza artificiale e responsabilità. Atti del Convegno, Università per Stranieri di Perugia, 29 novembre 2017*, Giappichelli, Torino, 2018, pp. 21–22

<sup>57</sup> Ivi, pp. 22–23.

<sup>58</sup>RUFFOLO, U. (a cura di), op. cit., pp. 23–25

<sup>59</sup> Ivi, p. 24.

assicurativi obbligatori, la cooperazione tra produttori e sviluppatori e, nei casi di rischio algoritmico sistemico, l'estensione della responsabilità oggettiva.<sup>60</sup>

Da un'altra prospettiva Annalisa Caligiuri, affronta il tema su un piano teorico e sistemico, evidenziando come l'intelligenza artificiale metta in crisi i presupposti concettuali del diritto classico strumenti concettuali, senza una riformulazione critica<sup>61</sup>. L'autrice individua nella crescente opacità decisionale dei sistemi intelligenti il principale fattore di crisi dell'imputazione soggettiva, e propone un superamento della logica binaria autore/danno attraverso l'elaborazione di modelli giuridici ibridi, ispirati alla responsabilità da attività pericolose o alla no-fault liability<sup>62</sup>.

Viene presa in esame anche l'ipotesi, ancora controversa, di attribuire una forma di personalità giuridica dell'algoritmo come entità autonoma ai fini della responsabilità, seppur con estrema cautela<sup>63</sup>. Più rilevante, tuttavia, è l'approccio funzionale, secondo cui chi ricopre un ruolo chiave nella progettazione o gestione del sistema deve essere gravato da obblighi giuridici specifici, anche in assenza di colpa.<sup>64</sup>

Si cercherebbe dunque di costruire una “rete di responsabilità” dinamica e adattiva, capace di riflettere la molteplicità di attori coinvolti nei processi automatizzati<sup>65</sup>. Il diritto è chiamato a un ruolo proattivo ed anticipatorio: non può più limitarsi a sanzionare a posteriori, ma deve guidare lo sviluppo tecnologico preservando i principi fondamentali dell'ordinamento<sup>66</sup>.

#### **4.5 Una tecnologia da orientare: il compito dell'etica tra sfide e responsabilità**

Il documento *AI4People*, frutto del lavoro multidisciplinare coordinato da Luciano Floridi, rappresenta uno dei tentativi più autorevoli di fornire una risposta sistematica e

---

<sup>60</sup> Ivi, p. 28.

<sup>61</sup> CALIGIURI, A. (a cura di), *Responsabilità e AI*, Giappichelli, Torino, 2022, p. 53

<sup>62</sup> Ivi, p.54.

<sup>63</sup> Ivi, p.55.

<sup>64</sup> Ivi, p.57.

<sup>65</sup> Ivi, p.58.

<sup>66</sup> CALIGIURI, A. (a cura di), op. cit., pp. 59–60.

valoriale alle sfide poste dall'intelligenza artificiale.<sup>67</sup> I cinque principi proposti individuati: beneficenza, non maleficenza, autonomia, giustizia ed esplicabilità, non sono solo mere astrazioni teoriche, ma si propongono come strumenti concreti ed operativi per orientare lo sviluppo e l'uso etico dei sistemi algoritmici.

La richiesta di spiegabilità, ad esempio, si intreccia direttamente con il principio di esplicabilità, inteso come intelligibilità e accountability nei processi decisionali automatizzati. La dignità umana, spesso minacciata dall'opacità e dalla standardizzazione algoritmica, trova tutela nella beneficenza, quale promozione del benessere e nella non maleficenza, ovvero nell'imperativo di evitare danni. La questione della discriminazione algoritmica richiama il principio di giustizia, ma anche quello di autonomia, poiché un sistema che penalizza arbitrariamente ed impedisce scelte libere e consapevoli. Il principio del controllo umano significativo, infine, si colloca all'incrocio tra autonomia, responsabilità ed esplicabilità, affermando il diritto dell'essere umano a comprendere, valutare e intervenire nel processo decisionale.

In quest'ottica, l'algoritmica non si limita a descrivere ciò che l'intelligenza artificiale non dovrebbe fare, ma propone una bussola normativa per definire ciò che dovrebbe diventare: una tecnologia integrata nei valori fondativi dello spazio giuridico e culturale europeo. Saranno proprio questi valori a costituire il criterio guida nell'analisi dei contributi offerti dai comitati etici internazionali e, successivamente, nella valutazione delle risposte giuridiche formalizzate nel GDPR e dall'AI ACT.

## **5. Etica e diritto in dialogo**

L'intelligenza artificiale non è più una prospettiva futura, ma una presenza pervasiva nei meccanismi che regolano la vita collettiva. Di fronte a questa trasformazione silenziosa ma profonda, è emersa con chiarezza l'esigenza di possibili interventi regolativi efficaci.

Negli ultimi anni, questa consapevolezza ha preso forma concreta attraverso una serie di iniziative promosse da comitati etici, gruppi di esperti ed organismi internazionali,

---

<sup>67</sup> FLORIDI, L., et al., *AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*, in *Minds and Machines*, vol. 28, 2018, pp. 689–707.

impegnati a cercare di tradurre i principi generali in strumenti operativi. Le linee guida, le raccomandazioni ed i codici etici prodotti da questi organismi non rappresentano delle semplici dichiarazioni d'intenti, ma costituiscono il primo nucleo di una governance dell'intelligenza artificiale, pensata per accompagnare l'innovazione con responsabilità.

E, proprio questi strumenti, oggi considerati parte di una “soft law” etica, sono destinati a costruire le fondamenta delle normative future, contribuendo a plasmare un diritto dell'intelligenza artificiale più maturo e consapevole.

Proprio queste esperienze, che segnano un passaggio decisivo dalla riflessione teorica all'azione concreta che si dedica la sezione che segue.

### **5.1 Il progetto *Ethically Aligned Design* dell'IEEE**

Tra le iniziative significative volte a promuovere un'etica condivisa nello sviluppo dell'intelligenza artificiale, un ruolo di primo piano è stato assunto dal progetto *Ethically Aligned Design* (EAD), promosso dall'IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Nato nel 2016, questo ambizioso programma ha preso forma attraverso un processo partecipativo su scala globale, coinvolgendo centinaia di esperti provenienti da ambiti disciplinari, culturali e geografici differenti.

Fin dall'inizio, l'obiettivo dichiarato è stato quello di guidare la progettazione dei sistemi autonomi e intelligenti alla luce dei valori fondamentali dell'essere umano.<sup>68</sup>

Un aspetto centrale dell'iniziativa è rappresentato dall'elaborazione di cinque principi generali che dovrebbero orientare l'intero ciclo di vita dell'IA: il rispetto dei diritti umani, la promozione del benessere individuale e collettivo, la responsabilità di progettisti e degli operatori, la trasparenza dei sistemi e la prevenzione di usi impropri o dannosi dell'IA.<sup>69</sup> A tali principi si affiancano indicazioni operative, come la necessità di

---

<sup>68</sup> IEEE GLOBAL INITIATIVE ON ETHICS OF AUTONOMOUS AND INTELLIGENT SYSTEMS, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, Version 1, IEEE, 2016, p. 3.

<sup>69</sup> Ivi, pp. 3–4.

implementare audit trail e sistemi di tracciabilità delle decisioni automatizzate, anche in presenza di tecnologie adattive o basate su apprendimento automatico.<sup>70</sup>

La seconda versione del documento (*EAD v2*) ha ampliato e approfondito ulteriormente questa impostazione, includendo sezioni dedicate alla governance dei dati personali, alla tutela dei diritti nel contesto digitale, alla responsabilità giuridica e alla definizione di nuove metriche per valutare il progresso tecnologico, tra cui il benessere umano, proposto come alternativa al tradizionale indicatore di produttività.<sup>71</sup>

Uno degli elementi di maggiore originalità del progetto è rappresentato dalla sua apertura alle diverse tradizioni etiche del mondo: accanto ai riferimenti all'etica classica occidentale -da Aristotele a Kant- si trovano citazioni ed ispirazioni provenienti da filosofie orientali e africane come il Confucianesimo, lo Shintoismo e l'etica Ubuntu. Questa scelta nasce dalla consapevolezza che un'intelligenza artificiale autenticamente al servizio dell'umanità deve essere culturalmente sensibile, inclusiva e capace di rispettare la pluralità dei contesti valoriali.<sup>72</sup>

*Ethically Aligned Design* ha costituito uno dei primi tentativi concreti di tradurre i principi etici in linee guida progettuali, contribuendo a definire standard internazionali e strumenti regolativi in grado di integrare l'etica nel cuore stesso del processo di innovazione. L'iniziativa ha così segnato un passaggio decisivo: da una visione dell'etica come elemento accessorio o postumo, a una concezione in cui essa si configura come architrave dell'intero sviluppo tecnologico, in dialogo costante con il diritto, la politica e la società civile.<sup>73</sup>

---

<sup>70</sup> Ivi, p. 159.

<sup>71</sup> IEEE GLOBAL INITIATIVE ON ETHICS OF AUTONOMOUS AND INTELLIGENT SYSTEMS, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, Version 2, IEEE, 2017, pp. 6–9 e 83–112.

<sup>72</sup> Ivi, pp. 2–3.

<sup>73</sup> IEEE GLOBAL INITIATIVE ON ETHICS OF AUTONOMOUS AND INTELLIGENT SYSTEMS, op. cit., pp. 3–5.

## 5.2 Il contributo dell'High-Level Expert Group on Artificial Intelligence

L'istituzione dell'High-Level Expert Group on Artificial Intelligence (AI HLEG), voluta dalla Commissione Europea nel 2018, ha rappresentato una tappa decisiva nell'evoluzione del quadro etico e giuridico dell'intelligenza artificiale all'interno dell'Unione Europea. Composto da 52 esperti provenienti da ambiti e formazioni totalmente eterogenee-università, mondo industriale, organizzazioni della società civile- il gruppo è stato incaricato di elaborare un quadro condiviso di valori e principi, capaci di orientare lo sviluppo dell'intelligenza artificiale verso forme affidabili, trasparenti e antropocentriche.<sup>74</sup>

Il risultato principale del loro lavoro è contenuto nelle *Ethics Guidelines for Trustworthy AI*, pubblicate nell'aprile 2019, in cui viene introdotto il concetto di **“Intelligenza artificiale affidabile”** (*Trustworthy AI*), fondata su tre pilastri inscindibili: rispetto del diritto, conformità e massima aderenza ai valori etici, robustezza sia tecnica che sociale<sup>75</sup>. Per dare concretezza a questi principi, il gruppo ha individuato sette requisiti chiave, tra cui la supervisione umana, la sicurezza, la trasparenza, la non discriminazione, e l'accountability- che fungono da criteri operativi per la progettazione e la valutazione dei sistemi intelligenti. Non si tratta di semplici enunciazioni astratte, ma di parametri funzionali a garantire che l'innovazione tecnologica si sviluppi in armonia con l'interesse collettivo.<sup>76</sup>

A completamento di queste linee guida, il gruppo ha poi pubblicato un documento successivo, *Policy and Investment Recommendations for Trustworthy AI*, che amplia l'analisi etica introducendo una serie di considerazioni e proposte strategiche e operative. Qui viene sottolineata l'urgenza di costruire ecosistemi sostenibili di innovazione, di rafforzare la cooperazione tra pubblico e privato, e di investire in percorsi formativi

---

<sup>74</sup>HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A Definition of AI: Main Capabilities and Scientific Disciplines*, European Commission, Bruxelles, 8 aprile 2019, p. 4

<sup>75</sup> *Ivi*, p. 7.

<sup>76</sup> HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *op. cit.*, pp. 12–16.

interdisciplinari, in grado di preparare professionisti e decisori alle sfide poste dall'adozione diffusa dell'AI<sup>77</sup>. Particolare attenzione è riservata all'adozione di strumenti di regolazione flessibili, come le *soft law*, e all'integrazione, fin dalle fasi iniziali dello sviluppo, di processi di valutazione dell'impatto etico sociale.<sup>78</sup>

Attraverso questi due testi, l'AI HLEG ha svolto un ruolo decisivo nell'impostare le basi concettuali e operative del futuro regolamento Europeo sull'IA (AI Act), dimostrando come un approccio basato sui valori possa tradursi in standard tecnici, linee guida politiche e strategie di investimento, capaci di collocare l'etica al centro della governance algoritmica.

L'analisi dei principali comitati etici europei — da *Ethically Aligned Design* all'azione propositiva di *AI4People*, fino alle linee guida dell'*High-Level Expert Group on AI* — ha evidenziato la volontà dell'Unione Europea di incardinare l'innovazione tecnologica all'interno di una cornice valoriale condivisa, ponendo l'etica come fondamento del processo regolativo. Tuttavia, proprio mentre si consolidava questa architettura teorica, è emersa una consapevolezza crescente: affinché l'etica non resti confinata a un piano dichiarativo, è necessario dotarsi di strumenti effettivi, operativi e vincolanti, capaci di tradurre i principi in norme giuridiche e scelte istituzionali concrete.

Si apre così una nuova fase nel dibattito sull'IA ed è in questo contesto che si collocano due documenti fondamentali: la *OECD Council Recommendation on Artificial Intelligence* del 2019, che rappresenta un impegno condiviso da parte dei paesi democratici e industrializzati, per promuovere un'AI responsabile e centrata sull'uomo; e la *Recommendation on the Ethics of Artificial Intelligence*, adottata dall'UNESCO nel 2021, che costituisce la prima cornice etica globale, fondata su diritti umani, sostenibilità ambientale e sul rispetto del pluralismo culturale.

Queste iniziative internazionali rappresentano un ponte tra l'etica e il diritto, un passaggio che prepara all'elaborazione di una regolazione giuridica matura, coerente e ambiziosa: il Regolamento europeo sull'AI (AI Act), prima proposta normativa vincolante.

---

<sup>77</sup> Ivi, pp. 5–9.

<sup>78</sup> Ivi, pp. 11–13.

### 5.3 OECD AI Principles (2019)

Nel maggio 2019, l'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE) ha adottato la prima raccomandazione intergovernativa sull'intelligenza artificiale, sottoscritta da oltre 40 Paesi, tra cui le principali economie occidentali. Questo documento rappresenta un momento di svolta nel dibattito globale sull'AI: per la prima volta, si tenta di armonizzare, su scala internazionale, una serie di principi condivisi per promuovere uno sviluppo tecnologico responsabile, umano-centrico e socialmente orientato al benessere collettivo.<sup>79</sup>

La raccomandazione si fonda su cinque principi fondamentali, che delineano un modello di intelligenza artificiale al servizio della persona e della società. Anzitutto, l'OCSE afferma la centralità del rispetto dello stato di diritto, dei diritti umani, di valori democratici e della diversità culturale, nella consapevolezza che ogni applicazione algoritmica debba sempre contribuire alla tutela della dignità umana. A ciò si aggiunge il principio di trasparenza, intesa come tracciabilità e spiegabilità dei processi algoritmici, così da consentire agli utenti ed i soggetti regolatori di comprendere e controllare il funzionamento dei sistemi. Centrale è anche la robustezza, che richiama la necessità di garantire sicurezza, resilienza e accuratezza tecnica, anche attraverso l'adozione di adeguati strumenti di gestione del rischio. Il principio di equità sottolinea l'urgenza di prevenire discriminazioni, bias sistemici o esclusioni, promuovendo inclusività e giustizia sociale. Infine, il principio di responsabilità impone lo sviluppo di adeguati meccanismi di governance, che individuino con chiarezza i soggetti responsabili delle decisioni automatizzate.<sup>80</sup>

Questi principi, pur non essendo giuridicamente vincolanti, vengono concepiti dall'OCSE come linee guida operative, da integrare in tutte le fasi del ciclo di vita dei sistemi di intelligenza artificiale, dalla progettazione iniziale alla distribuzione sul mercato, fino al monitoraggio e aggiornamento post-distribuzione. L'approccio promosso dall'OCSE è olistico, volto a coniugare innovazione tecnologica, sostenibilità sociale e tutela dei diritti,

---

<sup>79</sup> OECD, *Recommendation of the Council on Artificial Intelligence*, OECD Legal Instruments, OECD/LEGAL/0449, 22 maggio 2019, p. 3.

<sup>80</sup> OECD, op. cit., pp. 7–10

evitando al contempo le frammentazioni normative tra Paesi terzi e garantendo una cooperazione multilaterale.<sup>81</sup>

In una seconda fase, l'OCSE ha accompagnato la raccomandazione con indicazioni operative più dettagliate, volte a facilitare l'implementazione concreta dei principi enunciati. Tra le misure suggerite rientrano: l'adozione di sistemi di accountability basati su audit indipendenti, la valutazione preventiva degli impatti etici e sociali dei sistemi intelligenti, la costruzione di framework di co-creazione tra pubblico e privato, e lo sviluppo di infrastrutture digitali aperte. L'obiettivo dichiarato è quello di promuovere una cultura della responsabilità condivisa, in grado di anticipare i rischi, correggere eventuali criticità e distribuire in modo equo i benefici dell'innovazione algoritmica.<sup>82</sup>

#### **5.4 L'approccio etico dell'UNESCO all'intelligenza artificiale**

L'UNESCO ha assunto un ruolo essenziale nella definizione di un quadro etico globale per lo sviluppo e l'impiego dell'intelligenza artificiale. Nel novembre 2021, la Conferenza generale dell'Organizzazione ha approvato per acclamazione la *Raccomandazione sull'etica dell'intelligenza artificiale*, adottata all'unanimità da 193 Stati membri al termine di un articolato processo di consultazioni multilaterali e interdisciplinari durato oltre due anni<sup>83</sup>. Si tratta di un documento ambizioso, che si fonda su quattro valori fondamentali: la tutela dei diritti umani e della dignità, la promozione di società pacifiche e giuste, il rispetto della diversità culturale e l'adozione di un approccio sostenibile allo sviluppo tecnologico.

Questi valori si traducono in undici principi operativi, pensati per guidare in modo concreto la progettazione e l'adozione dell'IA. Tra i più rilevanti si segnalano: proporzionalità tra fini e mezzi, la non maleficenza, la trasparenza, la sicurezza dei sistemi, la responsabilità giuridica e sociale, la protezione dei dati personali e la

---

<sup>81</sup> Ivi, pp. 11–13

<sup>82</sup> OECD, *Policy Observatory: Implementing the OECD AI Principles*, OECD.AI, Paris, 2021, pp. 4–8.

<sup>83</sup> UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, adopted by the General Conference of UNESCO at its 41st session, Paris, 23 novembre 2021, p. 80.

supervisione umana attiva<sup>84</sup>. A differenza di altre linee guida a carattere meramente consultivo, la Raccomandazione dell'UNESCO è vincolante per gli Stati aderenti e si distingue per la sua efficacia operativa: ogni Stato aderente è infatti tenuto a redigere e presentare un rapporto di monitoraggio quadriennale all'Assemblea generale<sup>85</sup>. In questo modo, l'Organizzazione ha stabilito un meccanismo di verifica periodico, che rafforza il valore operativo del documento e ne assicura la concreta implementazione a livello nazionale.

Parallelamente, l'UNESCO ha pubblicato anche le *Linee guida per l'IA nell'educazione*, rivolte ai decisori politici. In esse vengono esaminate le implicazioni dell'intelligenza artificiale nei contesti scolastici e universitari, con l'obiettivo di promuovere un'educazione equa, inclusiva e consapevole, capace di preparare le nuove generazioni a un uso critico e responsabile delle tecnologie emergenti<sup>86</sup>.

L'accelerazione determinata dalla diffusione delle IA generative, come ChatGPT, ha tuttavia reso necessario un ulteriore aggiornamento tempestivo. Nel 2023, l'UNESCO ha pubblicato la guida *ChatGPT and Artificial Intelligence in Higher Education – Quick Guide*, seguita da un documento più ampio dedicato all'uso dell'IA generativa in ambito educativo e di ricerca.<sup>87</sup>

In queste nuove linee guida, l'Organizzazione richiama l'attenzione sui rischi derivanti da un utilizzo non regolamentato di tali strumenti generativi: dalla scarsa trasparenza dei modelli, all'opacità dei meccanismi decisionali fino alla diffusione di contenuti falsi o non verificabili (deepfake), allo sfruttamento improprio di dati personali e alla marginalizzazione di opinioni divergenti.<sup>88</sup>

Per affrontare queste criticità, l'UNESCO propone sette linee d'azione prioritarie rivolte ai governi. Tra esse figurano: l'adeguamento delle strategie nazionali sull'IA, l'introduzione di norme chiare sulla protezione dei dati (con esplicito riferimento al

---

<sup>84</sup> UNESCO, op. cit., p. 80.

<sup>85</sup> Ibidem.

<sup>86</sup> UNESCO, *AI and Education: Guidance for Policy-makers*, Paris, 2021, p. 81

<sup>87</sup> UNESCO, *ChatGPT and Artificial Intelligence in Higher Education – Quick Guide*, Paris, 2023.

<sup>88</sup> UNESCO, *Guidance for Generative AI in Education and Research*, cit., pp. 83–84.

GDPR), la regolamentazione del diritto d'autore per i contenuti generati, la formazione mirata per insegnanti e studenti, nonché la definizione di un'età minima per l'uso autonomo delle tecnologie generative.<sup>89</sup>

In particolare, Stefania Giannini, assistente al Direttore generale dell'UNESCO, ha denunciato pubblicamente l'incongruenza tra la velocità di adozione dell'IA generativa nelle aule scolastiche e l'assenza di adeguati strumenti di valutazione, sottolineando che “la IA è stata lanciata nella sfera pubblica senza discussione o revisione”<sup>90</sup>. In linea con questa preoccupazione, il documento raccomanda di fissare l'età minima a 13 anni per l'utilizzo di tali strumenti, garantendo sistemi di verifica dell'identità e consenso informato ed esplicito da parte dei genitori e tutori legali.<sup>91</sup>

Infine, nelle otto raccomandazioni operative per l'uso educativo delle IA generativa, l'UNESCO ribadisce con forza alcuni principi irrinunciabili: la tutela dell'autonomia decisionale umana, la promozione della pluralità delle opinioni, lo sviluppo di competenze critiche e tecniche, la validazione dei sistemi utilizzati e la garanzia di accessibilità culturale e linguistica degli strumenti digitali<sup>92</sup>. L'intelligenza artificiale, infatti, non deve mai sostituire la dimensione umana dell'apprendimento, ma può e deve diventare uno strumento prezioso per rafforzarla, a patto che sia governata con consapevolezza, prudenza e responsabilità.

## **6. Dall'etica al diritto: la svolta normativa dell'Unione Europea**

Il percorso fin qui tracciato attraverso il contributo di organismi internazionali, comitati etici e dichiarazioni valoriali ha reso evidente che il vero snodo per una regolazione efficace dell'intelligenza artificiale non può più essere affidato soltanto alla moral suasion dei principi etici. La crescente pervasività dell'IA nelle istituzioni pubbliche, nei mercati

---

<sup>89</sup> Ivi, pp. 84–85.

<sup>90</sup> UNESCO, *Generative AI and the Future of Education*, a cura di S. Giannini, Paris, 2023, p. 83.

<sup>91</sup> UNESCO, *Guidance for Generative AI in Education and Research*, cit., p. 84.

<sup>92</sup> UNESCO, *Guidance for Generative AI in Education and Research*, cit., p. 85.

e nella vita quotidiana impone un salto di paradigma, ovvero il superamento della fase della *soft law* e di compiere un salto di qualità: una regolazione giuridica dotata di efficacia vincolante e di strumenti sanzionatori adeguati.

In questo contesto, il diritto non si pone in alternativa all'etica, ma ne costituisce il naturale sviluppo. È proprio nel punto di incontro tra visione valoriale e tecnica giuridica che si apre lo spazio per una regolazione autenticamente umana dell'IA, capace di coniugare progresso tecnologico e salvaguardia dei diritti fondamentali e dunque di preservare l'umano all'interno di un mondo automatizzato. Una regolazione che, lungi dal limitarsi a rincorrere l'innovazione, si propone di guidarla consapevolmente, traducendo la complessità degli algoritmi in categorie giuridiche ed in regole capaci di garantire trasparenza, equità e giustizia. Questo passaggio dal piano etico a quello normativo segna una trasformazione nel modo in cui le democrazie intendono governare l'innovazione tecnologica.

In questo scenario, l'Unione Europea ha assunto un ruolo di leadership normativa a livello globale. Dopo anni di riflessioni etiche, la Commissione ha deciso di imboccare con decisione la via della regolazione *ex ante*, abbandonando un reattivo, basato sull'intervento successivo al verificarsi del danno, in favore di una valutazione preventiva del rischio. L'idea guida di questa impostazione è profondamente innovativa: l'intensità delle tutele giuridiche non deve essere uniforme, ma proporzionata al livello di rischio che un sistema IA comporta per i diritti fondamentali degli individui.

Tale principio costituisce l'asse portante dell'AI Act, il primo regolamento europeo sull'intelligenza artificiale. Come si vedrà nella trattazione che segue, esso rappresenta non solo una risposta giuridica ma a una realtà tecnologica in rapida evoluzione, ma soprattutto un tentativo ambizioso di costruire una governance algoritmica multilivello, fondata sull'integrazione armonica di etica, diritto, scienza e società.

## 6.1 AI ACT

Il Regolamento europeo sull'intelligenza artificiale, noto come **AI Act**, rappresenta il primo quadro normativo orizzontale a livello globale interamente dedicato alla disciplina dei sistemi di IA. Proposto dalla Commissione europea il 21 aprile 2021, il regolamento

nasce dall'esigenza di affrontare in modo sistematico le sfide giuridiche, sociali ed economiche poste dalla crescente diffusione di tecnologie algoritmiche in ambiti sia pubblici sia privati.<sup>93</sup>Dopo un lungo e articolato iter legislativo, il testo è stato approvato in via definitiva nel marzo 2024 e pubblicato nella Gazzetta Ufficiale dell'Unione europea il 12 luglio dello stesso anno, segnando il superamento definitivo della fase della soft law e delle mere raccomandazioni non vincolanti.<sup>94</sup>

Il regolamento è entrato formalmente in vigore venti giorni dopo la pubblicazione, ma ne è stata prevista un'applicazione scaglionata: le norme che vietano determinate pratiche ritenute inaccettabili saranno operative trascorsi sei mesi dall'entrata in vigore, mentre il corpus principale delle restanti disposizioni troverà applicazione dopo 24 mesi.<sup>95</sup>Questa dilazione temporale ha certamente una funzione strategica: consentire agli operatori economici e gli Stati membri di adeguarsi gradualmente al nuovo impianto normativo, evitando impatti bruschi sull'equilibrio del mercato e sullo sviluppo tecnologico.

L'AI Act si propone di tutelare i diritti fondamentali delle persone, garantendo che l'intelligenza artificiale venga sviluppata e utilizzata in modo etico, affidabile e conforme ai valori dell'Unione.<sup>96</sup>Parallelamente, mira a promuovere l'innovazione attraverso un quadro giuridico chiaro e prevedibile, che favorisca la fiducia degli utenti e degli investitori. La finalità complessiva è dunque duplice: da un lato, contenere i rischi associati all'uso dell'IA; dall'altro sostenere la competitività tecnologica e digitale dell'UE in un contesto sempre più globalizzato.<sup>97</sup>

Il regolamento assume anche una valenza geopolitica, poiché si applica a tutti i fornitori e utenti di sistemi di IA che operano nel mercato dell'Unione, anche se stabiliti in Paesi terzi. Tale estensione extraterritoriale esprime l'intenzione dell'UE di proporre uno

---

<sup>93</sup> CIRONE, E., *L'AI Act e l'obiettivo (mancato?) di promuovere uno standard globale per la tutela dei diritti fondamentali*, in *Quaderni AISDUE – Rivista quadrimestrale*, fasc. spec. n. 2/2024, ISSN 2975-2698, p. 1.

<sup>94</sup> Ibidem.

<sup>95</sup> Ibidem.

<sup>96</sup> CIRONE, E., op. cit., p. 2.

<sup>97</sup> Ibidem.

standard globale basato sulla centralità della persona umana e di affermare la propria leadership normativa nel contesto della governance tecnologica internazionale.<sup>98</sup>

## 6.2 Architettura normativa dell'AI Act e approccio basato sul rischio

L'AI Act adotta una struttura normativa innovativa, basata su un approccio graduale al rischio, che costituisce uno degli elementi originali e distintivi della strategia europea in materia di regolazione dell'intelligenza artificiale. L'obiettivo primario del regolamento è garantire che lo sviluppo e l'utilizzo dei sistemi di IA si svolgano nel pieno rispetto dei diritti fondamentali, della democrazia, dello Stato di diritto e della sicurezza pubblica<sup>99</sup>, senza tuttavia porre ostacoli alla crescita e alla competitività tecnologica europea (*considerando* 5, 17).

Il legislatore europeo ha individuato quattro livelli di rischio, graduati in funzione del potenziale impatto che ciascun sistema può avere sui diritti e le libertà degli individui:

1. **Rischio inaccettabile:** in questa categoria rientrano i sistemi IA ritenuti intrinsecamente lesivi dei valori fondamentali dell'Unione, e pertanto vietati in modo assoluto. Si tratta, ad esempio, di applicazioni che ricorrono a tecniche subliminali per distorcere ed alterare il comportamento di una persona, in modo da causarle o poterle causare danni fisici o psicologici, dei sistemi di social scoring da parte delle autorità pubbliche, e di alcune forme di riconoscimento biometrico remoto in tempo reale (*artt. 5 e 6; considerando* 22-25)<sup>100</sup>.
2. **Rischio elevato:** comprende quei sistemi che, per contesto di impiego e di finalità, possono incidere in modo significativo su diritti e libertà fondamentali. Essi sono elencati all'Allegato III e includono, tra gli altri, i sistemi di IA utilizzati in ambito occupazionale, sanitario, educativo, e per l'accesso a servizi essenziali. Tali sistemi sono sottoposti a **obblighi rigorosi**, come la valutazione di conformità ex ante, la registrazione nella banca dati europea, la documentazione tecnica, la

---

<sup>98</sup> Ivi, p. 3

<sup>99</sup> Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, 13 giugno 2024, relativo all'intelligenza artificiale, in GUUE L 168/1, considerando 5 e 17.

<sup>100</sup> Regolamento (UE) 2024/1689 (AI Act), op. cit., artt. 5 e 6; considerando 22–25.

garanzia della qualità dei dati, e l'adozione di misure di supervisione umana (*artt. 6-29; considerando 30-59*).<sup>101</sup>

3. **Rischio limitato:** riguarda applicazioni per cui è richiesto un obbligo di **trasparenza specifica**. È il caso di strumenti come i chatbot, i deepfake o sistemi di classificazione automatica, per i quali è necessario informare chiaramente l'utente che sta interagendo con una macchina (*art. 50; considerando 60-62*).<sup>102</sup>
4. **Rischio minimo o nullo:** include i sistemi di IA che presentano un impatto trascurabile sui diritti delle persone. In questo caso, non sono previsti obblighi normativi specifici. Appartengono a questa categoria i sistemi utilizzati nei videogiochi o nei filtri antispam (*considerando 64*).<sup>103</sup>

Per i sistemi ad alto rischio, l'AI Act introduce l'obbligo di una valutazione di conformità obbligatoria, che può essere svolta internamente dall'operatore o affidata a un organismo notificato, a seconda delle caratteristiche del sistema (*artt. 43-44*). In entrambi i casi, è prevista la registrazione nella banca dati pubblica prevista dall'art. 60.<sup>104</sup>

Sul piano istituzionale, il regolamento disegna un'articolata struttura di governance multilivello, fondata sulla cooperazione tra:

- la **Commissione europea**,
- le **autorità nazionali competenti**,
- e il nuovo **Ufficio europeo per l'intelligenza artificiale**, istituito con il compito di coordinare l'attuazione del regolamento a livello dell'Unione e di fornire assistenza agli Stati membri per garantire un'interpretazione uniforme delle disposizioni<sup>105</sup> (*artt. 63-68; considerando 81-88*).

Di rilievo anche il regime sanzionatorio previsto: in caso di violazioni gravi, come l'immissione sul mercato di sistemi vietati o la mancata osservanza degli obblighi per

---

<sup>101</sup> Ivi, art. 6 e Allegato III; considerando 30–59.

<sup>102</sup> Ivi, art. 50; considerando 60–62

<sup>103</sup> Ivi, considerando 64

<sup>104</sup> Ivi, artt. 43–44 e 60.

<sup>105</sup> Regolamento (UE) 2024/1689 (AI Act), artt. 63–68; considerando 81–88.

quelli ad alto rischio, il regolamento stabilisce sanzioni amministrative pecuniarie fino a 35 milioni di euro o al 7% del fatturato annuo globale dell'operatore <sup>106</sup>(art. 99).

Infine, l'AI Act riconosce specifici diritti agli utenti: tra questi il diritto ad essere informati quando interagiscono con un sistema IA, la possibilità di presentare reclami e di ottenere informazioni chiare e comprensibili sul funzionamento dei sistemi. Queste tutele si inseriscono nell'orizzonte dei principi di accountability e di protezione dei diritti fondamentali, che permeano l'intera architettura del regolamento. (art. 52; considerando 60-62, 71)<sup>107</sup>.

### **6.3 Le tensioni interne dell'AI Act: tra ambizione regolativa e fragilità strutturale**

Il Regolamento europeo sull'intelligenza artificiale nasce con l'ambizione di coniugare innovazione e protezione, presentandosi come il primo tentativo organico al mondo di disciplinare l'intelligenza artificiale in modo sistematico. Tuttavia, nonostante la sua struttura formale avanzata, l'AI Act presenta una serie di tensioni strutturali e valoriali, che mettono in discussione la sua coerenza interna e la sua effettiva capacità di governo dei fenomeni tecnologici complessi. Come evidenzia criticamente il Quaderno della Fondazione Astrid, il Regolamento rappresenta il risultato di un compromesso instabile tra esigenze eterogenee: un lato, la spinta a sostenere la competitività digitale europea; dall'altro, la necessità di evitare che l'uso dell'IA arrechi danni alla coesione sociale e all'equilibrio democratico ed ai diritti fondamentali.<sup>108</sup>

La prima criticità emerge già a livello concettuale. Il modello regolativo si basa su una classificazione dei sistemi di IA secondo un gradiente di rischio, criterio che richiama alla logica della precauzione e consente una modulazione degli obblighi giuridici. Questa struttura si scontra profondamente con la realtà dinamica e polimorfica dei sistemi algoritmici, che non sempre possono essere rigidamente incasellati secondo schemi predefiniti. La definizione stessa di "sistemi ad alto rischio" appare essere in più punti

---

<sup>106</sup> Ivi, art. 99.

<sup>107</sup> Ivi, art. 52; considerando 60-62 e 71

<sup>108</sup> FONDAZIONE ASTRID, *Quaderno n. 17 – L'AI Act: verso una regolazione europea dell'intelligenza artificiale*, estratto, p. 8.

astratta e indeterminata, generando margini interpretativi ampi e potenzialmente elusivi. In tal modo, il Regolamento rischia di essere al tempo stesso troppo ambizioso e troppo generico nelle prescrizioni, producendo un paradosso che ne indebolisce la portata applicativa.<sup>109</sup>

Ma i suoi grandi limiti non si esauriscono sul piano ingegneristico. Anche sotto il profilo assiologico, l'AI Act sembra adottare, almeno a livello dichiarativo, i principi ed i riferimenti, pilastri della cultura europea: dignità, autonomia, uguaglianza. Tuttavia, come sottolinea la stessa Fondazione Astrid, tali valori vengono evocati nei considerando senza essere tradotti in obblighi giuridici cogenti nella parte dispositiva. Il risultato è una tutela puramente nominale, che rischia di ridursi a retorica regolativa se non accompagnata da strumenti efficaci di garanzia.<sup>110</sup>

Particolarmente problematica risulta la gestione del rapporto tra IA e diritti fondamentali. L'attenzione normativa si concentra sui rischi legati alla salute, alla sicurezza e alle libertà individuali, ignorando o trattando in modo marginale altri aspetti cruciali come: la disuguaglianza sistemica prodotta da algoritmi opachi, il condizionamento comportamentale nei mercati digitali, o le derive discriminatorie nei processi decisionali automatizzati. In questo modo, il diritto positivo non solo rischia di restare indietro rispetto alla tecnologia, ma di legittimare, seppur involontariamente, modelli distorsivi di governance algoritmica.<sup>111</sup>

A tutto ciò si aggiunge un'ulteriore contraddizione: pur proponendosi come modello globale di regolazione, in grado di orientare anche altri ordinamenti, esso resta fortemente ancorato a una visione eurocentrica e autoreferenziale, incapace di dialogare in modo autentico con i pluralismi giuridico-culturali del contesto internazionale. Il risultato è uno strumento apparentemente “aperto” ma non del tutto “inclusivo”, con conseguenti difficoltà sia sul fronte dell'attuazione interna sia nel suo affermarsi come benchmark normativo internazionale.<sup>112</sup>

---

<sup>109</sup>Ivi, pp. 9–10.

<sup>110</sup> Ivi, pp. 10–11.

<sup>111</sup> FONDAZIONE ASTRID, op. cit., p. 12.

<sup>112</sup> Ivi, pp. 12–13.

È evidente come la criticità dell'AI ACT non riguardi solo il contenuto delle disposizioni, ma investa il modello di razionalità giuridica su cui esso si fonda. Come sottolinea il Quaderno Astrid, vi è il rischio che la regolazione dell'IA si trasformi in un esercizio tecnocratico di formalizzazione del rischio, privo di una base deliberativa autenticamente democratica. Una governance algoritmica che non si accompagni a una riflessione pubblica sui fini, sui valori e sulla giustizia finisce per svuotare la normazione della sua funzione critica e trasformativa.<sup>113</sup>

Un aspetto che merita ulteriore attenzione riguarda la scelta dell'AI Act di escludere dal proprio ambito di applicazione i sistemi di IA sviluppati o utilizzati esclusivamente per la ricerca scientifica (art. 2, par. 6). Se pur da un lato questa nasce dall'intento, certamente condivisibile, di tutelare la libertà accademica e non frenare il progresso dell'innovazione, dall'altra solleva interrogativi non trascurabili sul piano giuridico e pratico. Come osserva Di Tano, il riferimento normativo a sistemi sviluppati "al solo scopo" di ricerca e sviluppo scientifico risulta eccessivamente vago e privo di parametri oggettivi di verifica, aprendo la strada a margini interpretativi ampi e potenzialmente strumentalizzabili, soprattutto nei contesti di trasferimento tecnologico o di collaborazione tra enti pubblici e soggetti privati.<sup>114</sup>

Il rischio concreto è che progetti orientati, in realtà, a finalità operative o commerciali vengano presentati come iniziative di ricerca pura, approfittando di quella zona grigia normativa per eludere gli obblighi più stringenti previsti per i sistemi ad alto rischio. Tale ambiguità interpretativa può incentivare pratiche di *forum shopping* tra regimi giuridici più o meno favorevoli, compromettendo l'effettività della disciplina europea e creando una frattura tra l'intento regolativo e la sua reale implementazione<sup>115</sup>.

A questo si aggiunge una critica più ampia e strutturale: l'AI Act concentra la propria attenzione regolativa sulla fase applicativa dei sistemi di IA, trascurando però momenti altrettanto essenziali come la progettazione e l'addestramento. Eppure, è proprio in queste

---

<sup>113</sup> Ivi, p. 13.

<sup>114</sup>DI TANO, F., *Intelligenza Artificiale, ricerca scientifica e adeguamento etico-giuridico*, in *i-lex – Rivista di Scienze Giuridiche, Scienze Cognitive ed Intelligenza Artificiale*, vol. 18, n. 1, 2025, p. 21, DOI: 10.6092/issn.1825-1927/21064.

<sup>115</sup>Ivi, pp. 21–22.

fasi preliminari che si compiono molte delle scelte più rilevanti in termini di tutela dei diritti fondamentali. Ignorare questi momenti significa affidarsi a strumenti correttivi a posteriori, spesso inefficaci, quando gli effetti delle tecnologie sono già pienamente dispiegati e difficilmente reversibili. Come rilevato da parte della dottrina, si tratta di una forma di tutela “ritardata”, che giunge troppo tardi per svolgere una reale funzione preventiva<sup>116</sup>.

Sarebbe dunque auspicabile, come suggerisce la letteratura recente, prevedere l'introduzione di strumenti minimi di accountability anche nel campo della ricerca scientifica. Si pensi, ad esempio, a valutazioni d'impatto etico obbligatorie o a modelli di *ethics by design*, in grado di integrare i principi fondamentali già nella fase di concezione dei sistemi<sup>117</sup>. In assenza di tali garanzie, la ricerca scientifica rischia di trasformarsi in un vuoto normativo che potrebbe avere ripercussioni sia sul piano della protezione dei diritti sia sul principio di uniformità del diritto europeo<sup>118</sup>.

## **7. Le iniziative legislative italiane in materia di intelligenza artificiale: il D.d.l. n. 1146-B**

Un passaggio decisivo del percorso regolativo nazionale è rappresentato dal disegno di legge n. 1146-B, recante *Disposizioni e deleghe al Governo in materia di intelligenza artificiale*. Il provvedimento, approvato in prima lettura dal Senato il 20 marzo 2025 e successivamente modificato dalla Camera dei deputati il 25 giugno 2025 successivo, per poi tornare nuovamente al Senato il 26 giugno.<sup>119</sup>

Il testo si compone di 28 articoli ed è animato da un obiettivo chiaro: promuovere un utilizzo dell'intelligenza artificiale “corretto, trasparente e responsabile, sempre in una

---

<sup>116</sup> RESSEGUIER, A., UFERT, F., *AI research ethics is in its infancy: the EU's AI Act can make it a grown-up*, in *Research Ethics*, vol. 20, n. 2, 2024, p. 149.

<sup>117</sup> CASIRAGHI, S., VAN DIJK, N., *Ethics reviews in the European Union. Implications for the governance of scientific research in times of data science and Artificial Intelligence*, in *Law, Innovation and Technology*, vol. 16, n. 1, 2024, pp. 105–107.

<sup>118</sup> DI TANO, F., op. cit., pp. 22–23.

<sup>119</sup> D.d.l. n. 1146-B, *Disposizioni e deleghe al Governo in materia di intelligenza artificiale*, approvato dal Senato il 20 marzo 2025, modificato dalla Camera dei deputati il 25 giugno 2025 e nuovamente trasmesso al Senato il 26 giugno 2025.

prospettiva chiaramente antropocentrica.<sup>120</sup>L'idea di fondo è quella di inserire l'IA all'interno del tessuto sociale ed istituzionale senza snaturarne i valori di base ma anzi rafforzandoli in coerenza con i principi dell'Unione Europea e con l'impianto del Regolamento (UE) 2024/1689.

Di grande rilievo è L'art. 3, che delinea i principi generali destinati a guidare l'intero ciclo di vita dei sistemi di IA, dalla fase di ricerca e progettazione fino allo sviluppo e all'impiego concreto. Vengono individuati dunque i capisaldi della futura disciplina: trasparenza, sicurezza, tutela della dignità e non discriminazione, accuratezza e sostenibilità. Nel corso dell'esame parlamentare la Camera ha arricchito tale disposizione introducendo al comma 4 una clausola di garanzia dal forte significato politico e costituzionale: l'utilizzo dell'intelligenza artificiale non potrà mai comprimere lo svolgimento del metodo democratico, né condizionare indebitamente la libertà del dibattito politico. Contestualmente, si riafferma tanto la necessità di tutelare tanto la sovranità dello Stato quanto diritti fondamentali dei cittadini, collocandosi l'AI all'interno di una cornice che non è solo tecnica o giuridica, ma anche profondamente valoriale.<sup>121</sup>

Un secondo nucleo di disposizioni è dedicato all'architettura istituzionale della governance dell'intelligenza artificiale. L'art. 20 attribuisce all'Agenzia per l'Italia Digitale (AgID) e all'Agenzia per la Cybersicurezza Nazionale (ACN) il ruolo di Autorità nazionali competenti in materia di intelligenza artificiale, ciascuna con compiti distinti ma complementari: la prima orientata alla promozione e alla diffusione delle tecnologie digitali, la seconda focalizzata sulla protezione degli ecosistemi informatici e sulla prevenzione dei rischi connessi alla sicurezza. Accanto a queste trovano spazio le autorità di vigilanza settoriale Banca d'Italia, la CONSOB e l'IVASS, cui spetta la supervisione specifica dei sistemi di intelligenza artificiale classificati ad alto rischio in ambito finanziario, in conformità con l'art. 74, par. 6, AI Act.<sup>122</sup>

---

<sup>120</sup> Ivi, art. 1.

<sup>121</sup> D.d.l. n. 1146-B, *op. cit.*, art. 3, co. 4.

<sup>122</sup> Ivi, art. 20. *Art. 74, par. 6, AI Act, Reg. (UE) 2024/1689.*

A completare questo quadro, si colloca il Comitato di coordinamento, istituito dall'art 19, presieduto dal Presidente del Consiglio e composto da Ministri competenti insieme alle principali Autorità indipendenti. Tale organismo ha il compito di garantire una regia unitaria, fornendo indirizzo strategico e assicurando un raccordo interistituzionale in grado di evitare frammentazioni e sovrapposizioni.<sup>123</sup>

Come sottolineano Deriu e Racioppi, questa impostazione testimonia una chiara consapevolezza: la regolazione dell'IA non può ridursi ad un piano meramente tecnico-giuridico, ma deve assumere i tratti di una governance multilivello. Solo un modello capace di bilanciare le competenze delle diverse istituzioni e di coordinare l'azione nazionale con le dinamiche sovranazionali può infatti garantire un livello dell'AI efficace, coerente e in sintonia con la portata globale del fenomeno.<sup>124</sup>

Sul piano applicativo, il disegno di legge interviene a regolare l'utilizzo dell'intelligenza artificiale nei settori strategici. Particolare rilievo assume l'art. 14, dedicato alle pubbliche amministrazioni: qui si vede l'impiego dei sistemi di IA come strumenti per migliorare l'efficienza, ridurre i tempi dei procedimenti ed innalzare la qualità dei servizi resi ai cittadini. La norma, tuttavia è chiara nel fissare un limite invalicabile: l'algoritmo deve mantenere un ruolo meramente strumentale rispetto al potere decisionale umano. In questa prospettiva, viene ribadito che la responsabilità finale dei provvedimenti resta sempre in capo al funzionario pubblico. A ciò si aggiungono obblighi di tracciabilità e conoscibilità degli algoritmi impiegati, nonché misure specifiche di formazione per i loro utilizzatori, in modo da assicurare un uso consapevole e corretto delle tecnologie.<sup>125</sup> Come sottolineato nel Quaderno FinTech, tale disposizione si colloca in linea con i principi affermati dall'AI Act, ma tenta allo stesso tempo di rafforzare le garanzie a presidio della trasparenza amministrativa e della responsabilità democratica. Sullo sfondo si intravedono i principi giurisprudenziali già affermati: il diritto alla conoscibilità, il

---

<sup>123</sup> *Ivi*, art 19

<sup>124</sup> DERIU, P., RACIOPPI, S., *Riflessioni in tema di intelligenza artificiale e attività di vigilanza*, in *Quaderni FinTech*, n. 15, CONSOB, agosto 2025, p. 29.

<sup>125</sup> D.d.l. n. 1146-B, cit., art. 14.

divieto di esclusività della decisione algoritmica e il principio di non discriminazione, qui trasposti e consolidati all'interno di una cornice normativa nazionale.<sup>126</sup>

## **7.1 L'applicazione settoriale dell'intelligenza artificiale nel disegno di legge n. 1146-B**

Il d.d.l. 1146-B non si limita a fissare principi generali e regole organizzative, ma dedica un'attenzione particolare ai settori sensibili, ossia quei contesti nei quali l'intelligenza artificiale può incidere in maniera diretta e profonda sulla vita dei cittadini. Questa impostazione riflette una consapevolezza di fondo: il rischio connesso all'impiego dell'IA non è uniforme, ma varia a seconda degli ambiti applicativi. Per questo, accanto al modello europeo di regolazione trasversale, il legislatore nazionale ha scelto di individuare alcune aree prioritarie da presidiare con strumenti di garanzia specifici.

Un ruolo centrale è attribuito al settore della sanità e della disabilità. L'art. 7 sancisce che i sistemi di IA possono essere impiegati a supporto della ricerca clinica, della prevenzione e della diagnosi, garantendo sempre il rispetto dei principi di sicurezza, trasparenza e protezione dei dati personali. In questa logica si colloca anche l'art. 8, che introduce una disciplina mirata al riutilizzo dei dati sanitari per finalità di ricerca scientifica, prevedendo l'adozione di meccanismi di anonimizzazione e, ove possibile, l'utilizzo di dati sintetici, come misura ulteriore di tutela. L'art. 9 affida poi a un decreto legislativo del Governo il compito di definire regole semplificate per i trattamenti di dati destinati alla ricerca sanitaria, con l'obiettivo di bilanciare l'interesse pubblico al progresso scientifico con la salvaguardia della riservatezza. Infine, l'art. 10 istituisce una piattaforma nazionale di IA affidata all'AGENAS, chiamata ad esprimere i pareri preventivi sulla compatibilità dei sistemi con le esigenze di sicurezza dei pazienti e di protezione dei dati, collocandosi come snodo tecnico-istituzionale a presidio della correttezza e affidabilità delle innovazioni digitali in campo sanitario<sup>127</sup>.

Un ulteriore ambito di regolazione riguarda il **mondo del lavoro**. L'art. 11 del d.d.l. disciplina l'uso dei sistemi di IA nei rapporti di lavoro, imponendo ai datori di lavoro

---

<sup>126</sup> DERIU, P., RACIOPPI, S., op. cit., p. 31 ss.

<sup>127</sup> D.d.l. n. 1146-B, cit., artt. 7–10.

obblighi di trasparenza e informazione nei confronti dei lavoratori circa le modalità di utilizzo degli algoritmi nei processi di selezione, valutazione e gestione della prestazione. La norma stabilisce che tali strumenti non possono mai sostituire il potere decisionale umano, ma solo supportarlo, nel rispetto della dignità e dei diritti fondamentali del lavoratore. L'art. 12 istituisce inoltre un **Osservatorio nazionale sull'IA nel lavoro**, con funzioni di monitoraggio degli effetti occupazionali e di indirizzo delle politiche attive, a testimonianza della volontà del legislatore di governare l'impatto sociale della trasformazione tecnologica.<sup>128</sup>

Di particolare interesse è anche la disciplina dedicata alle **professioni intellettuali**. L'art. 13 consente l'impiego dell'IA da parte di avvocati, notai, medici, ingegneri e altri professionisti unicamente come strumento di supporto, escludendo la possibilità che l'algoritmo possa sostituire l'autonomia tecnica e decisionale del professionista. La norma impone inoltre l'obbligo di informare il cliente circa l'uso di sistemi di IA nel corso dell'attività professionale, rafforzando così il principio di trasparenza e la tutela dell'affidamento. La Commissione europea, in un parere circostanziato, aveva sollecitato il legislatore italiano a eliminare ogni restrizione ingiustificata all'uso dei sistemi non ad alto rischio, chiedendo una maggiore armonizzazione con la logica del Regolamento europeo; tale indicazione è stata in larga parte recepita nella versione definitiva del d.d.l.<sup>129</sup>

Un ulteriore snodo riguarda la **giustizia**. L'art. 15 stabilisce che i sistemi di IA possano essere impiegati nei procedimenti giudiziari soltanto a fini strumentali e di supporto, restando in ogni caso riservate al magistrato le decisioni su fatti, prove e diritto. L'obiettivo dichiarato è duplice: da un lato, incrementare l'efficienza e la celerità dei procedimenti; dall'altro, evitare che l'uso dell'IA possa indebolire i principi cardine del giusto processo, come il contraddittorio e la motivazione dei provvedimenti. Come

---

<sup>128</sup> Ivi, artt. 11–12.

<sup>129</sup>D.d.l. n. 1146-B, cit., art. 13; COMMISSIONE EUROPEA, *Parere circostanziato C(2024) 7814 sul d.d.l. italiano*, cit. in DERIU, P., RACIOPPI, S., *Riflessioni in tema di intelligenza artificiale e attività di vigilanza*, in *Quaderni FinTech*, n. 15, CONSOB, agosto 2025, p. 28.

osservato nel Quaderno FinTech, tale disposizione intende coniugare il potenziale innovativo degli strumenti predittivi con la salvaguardia dell'autonomia del giudice e della sovranità del diritto, ponendo l'accento sul carattere meramente ausiliario dell'algoritmo.<sup>130</sup>

Nel complesso, queste disposizioni settoriali confermano la volontà del legislatore italiano di integrare l'IA nei settori nevralgici dell'ordinamento senza snaturare i principi di fondo della tradizione giuridica nazionale ed europea. Sanità, lavoro, professioni e giustizia sono individuati come terreni nei quali l'uso dell'IA può apportare benefici significativi, ma richiede un rafforzamento delle garanzie. L'approccio seguito appare dunque improntato a un equilibrio: favorire l'innovazione, pur mantenendo salda la centralità della persona e la responsabilità degli operatori umani.

---

<sup>130</sup> Ivi, art. 15; DERIU, P., RACIOPPI, S., op. cit., p. 31.

## CAPITOLO 2

### Lavoro, intelligenza artificiale e tutela dei diritti fondamentali

#### 1. Trasformazioni del lavoro nell'era dell'intelligenza artificiale

##### 1.1 La trasformazione del lavoro nell'era digitale: oltre il determinismo tecnologico

Il dibattito sulla trasformazione dei lavori nell'era delle intelligenze artificiali si muove tra due poli completamente opposti. Da un lato, vi è chi immagina scenari distopici, segnati da una massiccia disoccupazione tecnologica e dall'obsolescenza del lavoro umano; Dall'altro vi è chi, affida alla digitalizzazione una funzione quasi salvifica, ritenendola capace di risolvere ogni contraddizione del presente. Entrambe queste prospettive, pur suggestive, appaiono riduttive e incapaci di restituire la complessità di questo nuovo fenomeno: complesso e ancora in divenire. Più realistico è considerare la digitalizzazione come un processo rapido e contraddittorio, ancora in corso, in cui il lavoro non scompare, ma si trasforma profondamente.<sup>131</sup>

In questo scenario, parlare di “Industria 4.0” non significa descrivere una realtà già pienamente compita, bensì richiamare una politica pubblica ed un percorso ancora in costruzione, che necessita di nuove visioni strategiche e di una forte capacità di governo. L'Italia, pur dimostrando una notevole vivacità creativa ed imprenditoriale, fatica però ad anticipare con decisione i cambiamenti strutturali imposti dall'innovazione tecnologica. Diventa essenziale alimentare un dibattito che si emancipi dalle letture semplicistiche, fondate sul determinismo tecnologico, secondo cui l'evoluzione della tecnica determinerebbe in modo automatica e unidirezionale i cambiamenti sociali ed economici. Sembra dunque necessario orientarsi verso una prospettiva di progettualità, che riconosca come la tecnica non sia mai neutrale e come il futuro del lavoro dipenderà, in ultima analisi, dalle scelte collettive e istituzionali che saremo in grado di compiere.<sup>132</sup>

---

<sup>131</sup> GRAMOLATI, A., CIPRIANI, A., MARI, G. (a cura di), *Il lavoro 4.0. La Quarta Rivoluzione industriale e le trasformazioni delle attività lavorative*, Firenze University Press, Firenze, 2018, p. XI.

<sup>132</sup> Ivi, pp. XI–XII.

Questa trasformazione non riguarda solo i processi produttivi, ma investe il significato stesso del lavoro. L'attività lavorativa tende progressivamente a superare la tradizionale contrapposizione tra manualità e intellettualità, aprendo la strada a una valorizzazione crescente delle dimensioni relazionali, cognitive e creative. In questo contesto prendono forma nuovi paradigmi, che pongono al centro la responsabilità, la libertà, la sostenibilità e soprattutto, la dignità della persona. Il lavoro si configura come uno spazio di espressione e autorealizzazione, fondato su qualità tipicamente umane come l'empatia, la capacità di negoziare, la collaborazione e il pensiero critico.<sup>133</sup>

Per affrontare questo passaggio epocale senza generare nuove fratture sociali, è necessario agire su più livelli: dalla promozione di una formazione continua alla redistribuzione del tempo di lavoro, dalla creazione di nuove filiere occupazionali fino alla protezione di coloro che possiedono competenze inadeguate sul piano digitale e rischiano l'esclusione dal mercato digitalizzato. Solo riaffermando con forza il primato della persona sulla macchina sarà possibile orientare l'innovazione verso il benessere del singolo e della collettività nella direzione di uno sviluppo autenticamente umano.<sup>134</sup>

In questa cornice, il riferimento all'"Industria 4.0" assume un significato particolare: non soltanto come espressione simbolica del cambiamento, ma come paradigma capace di orientare politiche pubbliche, strategie imprenditoriali e modelli organizzativi.

## **1.2 Il paradigma 4.0: una trasformazione tecnologica, sociale e simbolica**

L'espressione "Industria 4.0" è diventata, nell'arco di poco più di un decennio, come un'etichetta concettuale onnipresente, utilizzata per descrivere la profonda trasformazione dell'apparato produttivo industriale in chiave digitale, automatizzato e interconnessa. Tuttavia, dietro la sua apparente chiarezza, il concetto rivela contorni sfumati, aperti a interpretazioni diverse e talvolta persino contraddittorie. Nato ufficialmente in occasione dell'Hannover Messe del 2011, il termine "Industrie 4.0" fu inizialmente utilizzato nel contesto tedesco per designare un progetto di modernizzazione

---

<sup>133</sup> GRAMOLATI, A., CIPRIANI, A., MARI, G. (a cura di), op. cit., pp. XIII–XIV

<sup>134</sup>Ivi, pp. XIV–XV.

tecnologica dei sistemi produttivi, basato sull'integrazione sistematica dell'internet delle cose e dei sistemi cyber-fisici nei processi industriali.<sup>135</sup>

Successivamente, la declinazione statunitense ha introdotto il concetto di "Industrial Internet", ponendo l'accento sull'interconnessione tra macchine, impianti e dati attraverso la rete<sup>136</sup>. Tuttavia, al di là delle specificità geografiche, è l'intero impianto concettuale sotteso al paradigma 4.0 a mostrare una complessità strutturale che merita un'analisi attenta ed approfondita.

Nel dibattito contemporaneo, il termine "4.0" sembra aver assunto una funzione più performativa piuttosto che meramente descrittiva: non si limita a registrare un cambiamento oggettivo, ma contribuisce esso stesso a costruire una nuova visione della produzione e del lavoro, orientando politiche pubbliche, strategie aziendali e percorsi formativi. In questo senso, il "4.0" va inteso come una vera e propria grammatica della trasformazione, che non solo interpreta, ma al tempo stesso plasma l'evoluzione dei contesti socio-produttivi. Rimane tuttavia, come osserva Seghezzi, una significativa ambiguità semantica: il concetto di Industria 4.0 viene spesso utilizzato per descrivere realtà anche molto diverse tra loro, rischiando di diventare una formula vuota, più utile a sostenere narrazioni suggestive che a comprendere a fondo i fenomeni concreti.<sup>137</sup>

L'elemento che sembra unificare le diverse interpretazioni è l'idea di un'"intelligenza distribuita", che non riguarda più esclusivamente la componente umana del lavoro, ma si estende anche alle macchine, ai processi e agli algoritmi. Questo passaggio determina una duplice ridefinizione: da un lato, dell'organizzazione produttiva, che diventa progressivamente più orizzontale, flessibile e adattiva; dall'altro, del ruolo stesso del lavoratore, chiamato a collocarsi all'interno di una rete di connessioni digitali in cui competenze, responsabilità e margini di autonomia vengono continuamente rinegoziati. La fabbrica 4.0 si configura allora come un "ecosistema intelligente", all'interno del quale

---

<sup>135</sup>SEGHEZZI, F., *Tecnologie e lavoro: l'Italia e la quarta rivoluzione industriale*, in *Quaderni di Adapt*, 2020, p. 21

<sup>136</sup>Ivi, p. 22.

<sup>137</sup>Ivi, p. 23.

l'interazione tra uomo e macchina non si esaurisce nella dimensione puramente strumentale, ma assume anche un significato relazionale e cognitivo.<sup>138</sup>

Tuttavia, questo processo non è affatto neutro. Al contrario, comporta profonde implicazioni di natura economica, giuridica ed etica. La digitalizzazione integrata dei processi produttivi introduce nuove asimmetrie di potere tra capitale e lavoro, accentua le dinamiche di precarizzazione, rende obsoleti interi segmenti professionali e impone una ridefinizione urgente delle tutele, tanto collettive quanto individuali. Come sottolinea ancora Seghezzi, il rischio è che la narrazione ottimistica del 4.0 finisca per agire come una copertura ideologica ad operazioni di ristrutturazione che, sotto l'apparenza dell'innovazione, perseguono in realtà obiettivi di mera riduzione dei costi e intensificazione dei ritmi produttivi.<sup>139</sup>

In definitiva, parlare di “4.0” significa confrontarsi con un concetto stratificato, che racchiude in sé al tempo stesso promesse di efficienza e crescita irrisolte, insieme a tensioni latenti. Solo un approccio critico e multidisciplinare, capace di intrecciare l'analisi tecnologica con la riflessione sociologica e la valutazione giuridica, può restituire la complessità reale delle trasformazioni in atto, evitando di ridurle a semplici slogan.

Per comprendere la portata effettiva di questa trasformazione, è utile osservare in quali settori l'intelligenza artificiale stia trovando maggiore applicazione, verificando come la narrazione del “4.0” si traduca in pratiche concrete nei diversi comparti produttivi.

### **1.3 I settori economici più inclini all'adozione dell'IA**

L'adozione dell'intelligenza artificiale sta progressivamente interessando quasi tutti i settori economici, coinvolgendo non soltanto i comparti più tecnologicamente avanzati, ma anche quelli tradizionalmente meno inclini all'innovazione. Questo processo trasversale non è spinto unicamente dall'esigenza di sollevare la forza lavoro da mansioni ripetitive e poco qualificanti, ma risponde anche alla volontà di accrescere l'efficienza dei servizi e dei processi aziendali.

---

<sup>138</sup> SEGHEZZI, F., op. cit., pp. 24–25.

<sup>139</sup> Ivi, p. 27.

Secondo l'indagine *The Global Artificial Intelligence Trends 2020*, condotta da Analytics Insight su un campione di oltre 2.000 professionisti operanti in vari settori e aree geografiche, è l'industria ad alta tecnologia a guidare il cambiamento, con un tasso di adozione pari al 37%. Seguono il settore retail (9,1%) e quello sanitario e farmaceutico (8%). I comparti più tradizionali, come le infrastrutture e il turismo, mostrano invece una maggiore resistenza, pur manifestando una crescente apertura rispetto al passato. L'impiego dell'IA risulta particolarmente rilevante nella ricerca e sviluppo (34,3%), nei servizi finanziari (18,2%) e nel customer service (9,5%). Tra le priorità delle imprese che adottano queste tecnologie spiccano l'automazione dei processi (18,3%) e l'iperpersonalizzazione dell'offerta (7,8%). Restano, tuttavia, significativi ostacoli alla piena integrazione dell'IA: la carenza di competenze (31,1%), la limitata disponibilità di dati di qualità (27,7%) e gli elevati costi di implementazione (22,7%).<sup>140</sup>

A conferma e integrazione di tali evidenze, un'analisi più recente condotta dall'OCSE (*Artificial Intelligence and the Changing Demand for Skills in the Labour Market*, 2024) adotta un approccio diverso, basato sulla misurazione dell'"esposizione" delle occupazioni ai progressi dell'IA. Lo studio prende in esame oltre 700 professioni e l'universo delle inserzioni di lavoro in dieci Paesi OCSE, utilizzando un indicatore che mette in relazione i compiti caratteristici di ciascuna occupazione con le capacità delle applicazioni di IA. I risultati mostrano che circa un terzo delle vacancy si concentra in occupazioni ad alta esposizione – generalmente qualificate e ad elevata retribuzione, come consulenti genetici, esaminatori finanziari e attuari – mentre i comparti meno esposti coincidono con quelli a prevalente contenuto fisico non di routine, come i lavoratori dell'edilizia o gli istruttori sportivi. L'indagine rileva inoltre che, nelle occupazioni più esposte, la domanda di competenze manageriali e di business (gestione di progetti, finanza, amministrazione e compiti clericali) ha registrato una crescita costante nel tempo, accompagnata da un aumento di circa otto punti percentuali nella richiesta di competenze cognitive, digitali ed emotive. Tuttavia, un'analisi a livello di stabilimento, volta a isolare con maggiore precisione l'effetto dell'IA da altri fattori macroeconomici, segnala che proprio in tali contesti la domanda per queste competenze

---

<sup>140</sup> ANALYTICS INSIGHT, *The Global Artificial Intelligence Trends 2020*, cit. in I-COM, *L'intelligenza artificiale nei diversi settori: dove siamo e dove stiamo andando*, 2020, disponibile su: <https://www.i-com.it/2020/11/26/intelligenza-artificiale-settori/>.

inizia a ridursi, suggerendo una tendenza alla sostituzione delle skill più standardizzabili e routinarie. Lo studio OCSE offre così una prospettiva complementare alle rilevazioni campionarie, evidenziando non solo la diversa penetrazione settoriale dell'IA, ma anche i mutamenti qualitativi nella domanda di competenze, che assumono un rilievo cruciale per comprendere gli effetti di lungo periodo sul mercato del lavoro.<sup>141</sup>

Tuttavia, sul piano etico va rivelato che l'impatto dell'intelligenza artificiale sul lavoro non si misura soltanto in termini di produttività ed efficienza. Un nodo cruciale riguarda l'inclusione sociale, in particolare delle persone con disabilità, che potrebbero trarre nuove opportunità dalla digitalizzazione ma che rischiano anche di subire forme inedite di esclusione.

## **2. Lavoro, disabilità e intelligenza artificiale: opportunità di inclusione e rischi di esclusione sociale**

### **2.1 Inclusione lavorativa delle persone con disabilità**

L'intelligenza artificiale sta emergendo come uno strumento strategico capace di aprire nuove prospettive di inclusione lavorativa per le persone con disabilità, offrendo modalità innovative di accesso, apprendimento e interazione nei contesti aziendali. I dati del *Workmonitor Pulse* condotto da Randstad su un campione di 800 lavoratori italiani confermano questa tendenza: oltre la metà degli intervistati con disabilità (51%) ritiene che l'AI faciliti lo svolgimento delle proprie mansioni, mentre il 63% dichiara di utilizzarla almeno una volta a settimana, una percentuale significativamente più alta rispetto ai colleghi senza disabilità (37%).<sup>142</sup>

Il dato è ancora più rilevante se si considera la percezione soggettiva: più della metà delle persone con disabilità intervistate percepisce l'AI come uno strumento semplice da apprendere e capace di rendere il lavoro più stimolante, dimostrando come la tecnologia non sia solo un ausilio tecnico, ma anche un fattore di motivazione e crescita

---

<sup>141</sup> GREEN, A., *Artificial Intelligence and the Changing Demand for Skills in the Labour Market*, OECD Artificial Intelligence Papers, n. 14, OECD Publishing, Paris, aprile 2024.

<sup>142</sup> RANDSTAD, *Workmonitor Pulse – Speciale Intelligenza Artificiale e Inclusione*, 2024, disponibile su: <https://www.randstad.it>, ultimo accesso: 4 agosto 2025.

professionale. Inoltre, il 57% considera l'intelligenza artificiale un supporto efficace e complementare alla supervisione umana nel ridurre le disuguaglianze, mentre il 52% ne riconosce il potenziale nel rafforzare l'equità all'interno delle dinamiche aziendali.

Un altro dato significativo riguarda la formazione: il 56% dei lavoratori con disabilità ha ricevuto un training specifico sull'uso dell'AI, contro il 35% dei colleghi. Questa differenza segnala come molte imprese italiane stiano orientando le politiche di formazione proprio verso i profili più vulnerabili, riconoscendo nell'AI un alleato per accrescere le possibilità di inclusione e ridurre i gap preesistenti.<sup>143</sup>

Oltre ai dati empirici, anche la riflessione teorica più recente ha evidenziato come l'intelligenza artificiale possa rappresentare un'occasione per ripensare l'inclusione lavorativa a partire da una dimensione più ampia, legata alla libertà e alla partecipazione del lavoratore. In particolare, Massimiliano Costa invita a leggere il lavoro nell'ecosistema digitale e robotico attraverso due categorie fondamentali: la *capability to choice*, intesa come possibilità concreta di compiere scelte autonome rispetto al proprio percorso professionale, e la *capability to voice*, intesa come diritto ad esprimere la propria opinione e di incidere sui processi organizzativi. Entrambe queste dimensioni, se sostenute da politiche formative inclusive e da un uso eticamente orientato della tecnologia, possono rafforzare il senso di autodeterminazione, partecipazione e dignità lavorativa, specialmente per coloro che, come le persone con disabilità, si trovano spesso ai margini dei processi decisionali e produttivi. In tal senso, l'intelligenza artificiale non dovrebbe essere considerata soltanto come uno strumento tecnico, ma come un potenziale facilitatore di giustizia organizzativa, a condizione che venga inserita in un modello di lavoro, capace di riconoscere ed affrontare le disuguaglianze, valorizzando la soggettività di ciascun individuo.<sup>144</sup>

Un esempio concreto di innovazione inclusiva è rappresentato da **Maps for Easy Paths (MEP)**, un sistema di navigazione collaborativo progettato per persone con disabilità motorie. A differenza degli approcci tradizionali, fondati su rilevazioni statiche, MEP

---

<sup>143</sup> RANDSTAD, op. cit., 2024.

<sup>144</sup> COSTA, M., *Il lavoro nella transizione digitale e robotica: le dimensioni della capability to choice e quella to voice per la libertà realizzativa e di scelta del lavoratore*, in *MeTis. Mondi educativi. Temi, indagini, suggestioni*, vol. 14, n. 1, 2024, pp. 1–16.

ricostruisce i percorsi realmente accessibili grazie ai tracciati generati dagli utenti tramite i sensori dei dispositivi mobili. I dati raccolti vengono elaborati in cloud, corretti con algoritmi di machine learning e map-matching su OpenStreetMap, così da fornire percorsi aggiornati e affidabili. L'app, gratuita, consente sia la navigazione evitando ostacoli sia la segnalazione diretta di barriere urbane (gradini, rampe, parcheggi selvaggi, lavori in corso). Questo modello non solo riduce i costi di rilievi statici, ma trasforma l'esperienza individuale in un **bene comune**, rendendo la città più accessibile e inclusiva.<sup>145</sup>

A completamento di questa visione, la letteratura pedagogica più recente ha sottolineato l'importanza di valorizzare l'esperienza incarnata e la narrazione autobiografica come strumenti decisivi di empowerment per le persone con disabilità. In particolare, le ricerche di D'Angelo, Giacconi e collaboratori hanno messo in evidenza come lo sport e le pratiche di autorappresentazione possano diventare occasioni decisive di rinascita identitaria, favorendo una ri-significazione del corpo e del sé. Le biografie di atleti con disabilità testimoniano percorsi di riappropriazione del proprio spazio sociale e relazionale, nei quali l'agire corporeo si intreccia con la costruzione di senso, appartenenza e dignità. Questo approccio esperienziale non solo riconosce la soggettività delle persone con disabilità, ma la pone al centro di un processo trasformativo che si lega strettamente alle potenzialità dell'intelligenza artificiale. Se pensata in chiave inclusiva, infatti, l'AI può andare oltre la sua funzione strumentale, diventando un supporto capace di riconoscere e sostenere la persona nella sua unicità, rafforzando tanto la *capability to choice* quanto la *capability to voice*. Ne deriva che l'inclusione lavorativa non può limitarsi a un piano tecnico e organizzativo, ma deve radicarsi un'attenzione pedagogica capace di mettere al centro la persona, la sua storia e la sua corporeità. Solo così la tecnologia può trasformarsi in una leva autentica di autodeterminazione e dignità, contribuendo ad un modello di lavoro realmente inclusivo e rispettoso della pluralità delle esperienze umane.<sup>146</sup>

---

<sup>145</sup> AGENDA DIGITALE, *Un aiuto dalla tecnologia per muoversi in città: Maps for Easy Paths (MEP)*, s.d., disponibile su: <https://www.agendadigitale.eu>, ultimo accesso: 1° settembre 2025.

<sup>146</sup> KRAUS, C., GIACONI, C., D'ANGELO, A., *Different bodies: ibridazioni e trasformazioni nelle storie biografiche di persone con disabilità*, in *MeTis. Mondi educativi. Temi, indagini, suggestioni*, vol. 12, n. 2, 2022, pp. 19–32.

## 2.2 Il digital divide come barriera invisibile all'inclusione

Se da un lato l'intelligenza artificiale sembra aprire scenari di maggiore accessibilità e partecipazione, dall'altro non si può trascurare la questione delle condizioni che ne permettono un utilizzo realmente equo. Le persone con disabilità, infatti, non corrono il rischio di essere escluse soltanto per barriere di tipo fisico o cognitivo, ma anche per una frattura più sottile e diffusa: quella digitale. È qui che assume una grande centralità il concetto di *digital divide*, oggi sempre più complesso e articolato nelle sue implicazioni.

Non è più sufficiente interpretare il *digital divide* come una semplice separazione tra chi ha accesso e chi non ha accesso alle tecnologie digitali. Oggi, si tratta, di un divario multidimensionale, che coinvolge aspetti culturali, educativi, relazionali e persino simbolici. La questione non riguarda unicamente la disponibilità di dispositivi o connessioni, ma tocca in profondità la possibilità stessa di partecipare alla vita sociale e culturale in una società sempre più digitalizzata.

Il digital divide, si manifesta come esclusione da spazi di comunicazione, conoscenza e relazione, andando ad intrecciarsi con i processi di costruzione dell'identità e dell'appartenenza. Non basta, dunque, colmare i vari gap infrastrutturali: occorre interrogarsi su chi produce le tecnologie, quali usi esse promuovono e, soprattutto, quali soggetti vengono marginalizzati o resi invisibili nei nuovi sistemi digitali. Il digital divide si configura non soltanto come un problema tecnico, ma come una sfida pedagogica e politica, che richiede uno sguardo critico e trasformativo. Affrontarlo significa ripensare le condizioni di accesso e i criteri di progettazione delle tecnologie, per garantire che l'innovazione digitale non diventi una nuova forma di esclusione, ma un reale strumento di giustizia e inclusione sociale.<sup>147</sup>

All'interno di questa cornice, la condizione delle persone con disabilità evidenzia in modo particolarmente chiaro la natura complessa e stratificata del digital divide. Come osservano Kraus, Giacconi e D'Angelo, le barriere digitali non si esauriscono nella mera

---

<sup>147</sup> KRAUS, C., GIACONI, C., D'ANGELO, A., *Abitare le differenze: il digital divide nelle disabilità*, in *MeTis. Mondi educativi. Temi, indagini, suggestioni*, vol. 12, n. 2, 2022, pp. 57–58.

mancanza di strumenti o di connessioni, ma si radicano anche nelle modalità stesse con cui il digitale viene pensato e progettato. Troppo spesso, infatti, esso si fonda su un *modello implicito di normalità* che esclude corpi, menti e linguaggi non conformi<sup>148</sup>

Qui occorre introdurre il concetto di **etica by design**, che sottolinea come le tecnologie non siano mai neutre, ma incorporino scelte normative, sociali ed etiche già nella fase progettuale. Secondo Devon e van de Poel, il design è infatti un processo sociale, in cui non conta soltanto il prodotto finale, ma anche le modalità con cui vengono prese le decisioni: chi partecipa, con quali informazioni, con quali rapporti di potere e con quali possibilità di revisione delle scelte. Progettare “by design” significa dunque strutturare il processo in modo esplicito, inclusivo e responsabile, affinché l’innovazione non si traduca in nuove forme di esclusione, ma in strumenti di giustizia e sostenibilità. In questa prospettiva, il digital divide appare come uno degli esempi più chiari di come un design non inclusivo possa generare barriere sociali. Esso si manifesta come esclusione da spazi di comunicazione, conoscenza e relazione, intrecciandosi con i processi di costruzione dell’identità e dell’appartenenza. Non è sufficiente, quindi, colmare i gap infrastrutturali: occorre chiedersi chi produce le tecnologie, quali usi esse promuovono e soprattutto quali soggetti vengono marginalizzati o resi invisibili nei nuovi sistemi digitali. Il digital divide non è soltanto un problema tecnico, ma una sfida pedagogica e politica, che richiede uno sguardo critico e trasformativo. Affrontarlo significa ripensare le condizioni di accesso e i criteri di progettazione delle tecnologie, affinché l’innovazione digitale diventi un reale strumento di inclusione sociale.<sup>149</sup>

In questa prospettiva, la disabilità si configura come una sorta di cartina di tornasole delle dinamiche di esclusione, poiché mette in crisi l’idea di una presunta universalità dei dispositivi e dei codici digitali.<sup>150</sup> Ciò che si presenta come neutrale e valido per tutti, in realtà tradisce una visione parziale ed omologata, incapace di riconoscere la pluralità delle esperienze umane. Il divario digitale diventa a pieno titolo una questione politica ed epistemica: riguarda non solo l’accesso agli strumenti, ma la possibilità stessa di

---

<sup>148</sup>KRAUS, C., GIACONI, C., D’ANGELO, A., op. cit., p. 58.

<sup>149</sup>DEVON, R., VAN DE POEL, I., *Design Ethics: The Social Ethics Paradigm*, in *International Journal of Engineering Education*, vol. 20, n. 3, 2004, pp. 461–469.

<sup>150</sup>KRAUS, C., GIACONI, C., D’ANGELO, A., op. cit., p. 58.

legittimare il sapere, la voce e la presenza delle soggettività non standardizzate nello spazio digitale.<sup>151</sup>

Questo tipo di esclusione non si limita a produrre disuguaglianze di accesso, ma contribuisce a consolidare l'idea che esistano soggetti "fuori posto" nello spazio digitale, alimentando l'invisibilità tanto simbolica quanto materiale. Come osservano gli autori, ciò che viene negato non è solo l'accesso agli strumenti, ma «la possibilità stessa di esistere digitalmente».<sup>152</sup>

Non è più sufficiente limitarsi a denunciare l'esistenza del digital divide: occorre compiere un passo ulteriore e trasformare la questione dell'accesso digitale in una vera e propria politica di inclusione. Come ricorda l'UNESCO, l'inclusione digitale non si esaurisce nella semplice disponibilità di dispositivi o di connettività, ma implica la costruzione di un ambiente digitale progettato per valorizzare le diversità, favorire la partecipazione attiva e garantire a tutti la possibilità di contribuire alla società digitale.<sup>153</sup>

In questo senso, il concetto di *digital inclusion* si configura come un processo intenzionale e sistemico che mira non solo a rimuovere barriere tecniche, ma anche a sostenere le competenze e a promuovere l'autonomia dei soggetti, in particolare di quelli più vulnerabili. L'obiettivo non è solo “*portare dentro*” chi è rimasto ai margini, ma ridefinire gli spazi digitali in modo che risultino realmente condivisi, accessibili e giusti<sup>154</sup>.

Le difficoltà di accesso alle tecnologie si intrecciano con un'altra forma di disuguaglianza, meno visibile ma altrettanto incisiva: quella prodotta dai bias algoritmici. Essi rappresentano una barriera sottile e spesso invisibile, capace di incidere direttamente sulla parità di trattamento nel mercato del lavoro.

---

<sup>151</sup> KRAUS, C, GIACONI, C, D'ANGELO, A., *op. cit.*, p.60

<sup>152</sup> Ivi, p. 61.

<sup>153</sup> UNESCO, *Moving from digital divide to digital inclusion: key principles for digital inclusion*, Paris, 2022, p. 3

<sup>154</sup> Ivi, p. 6.

### 3. Disuguaglianze Digitali: I Bias Algoritmici nel Mercato del Lavoro

#### 3.1 Bias algoritmici e discriminazioni nei processi decisionali

L'impiego sempre più pervasivo di sistemi di intelligenza artificiale e di processi decisionali automatizzati ha reso il tema dei *bias algoritmici* una delle questioni centrali del dibattito giuridico ed etico contemporaneo. Con questa espressione si intendono le distorsioni sistematiche che si manifestano negli output degli algoritmi e che possono generare effetti discriminatori nei confronti di individui o gruppi, anche in assenza di un intento lesivo da parte di un soggetto umano<sup>155</sup>.

Le cause di tali distorsioni sono molteplici: possono derivare da un'errata impostazione delle regole in fase di programmazione, dalla selezione o dalla scarsa qualità dei dati di addestramento, oppure dall'impiego di variabili *proxy* che, pur non direttamente collegate a caratteristiche protette, finiscono per rappresentare un indicatore implicito. Nei sistemi *rule-based*, fondati su regole fisse, il rischio principale è legato alle scelte compiute dal programmatore; nei sistemi di *machine learning*, invece, l'algoritmo apprende in modo automatico dai dati, con la conseguenza di riprodurre e in alcuni casi persino amplificare, le disuguaglianze già presenti<sup>156</sup>.

La peculiarità della discriminazione algoritmica, rispetto a quella tradizionale di matrice umana, risiede proprio nell'assenza di intenzionalità: il trattamento differenziato può scaturire esclusivamente dall'interazione fra dati e modello matematico, senza che vi sia un intento consapevole di ledere. Questo aspetto la rende più difficile da individuare, contestare e provare, soprattutto a causa dell'opacità e della complessità tecnica che contraddistingue molti sistemi di intelligenza artificiale<sup>157</sup>.

Il fenomeno può assumere forme differenti: dalla discriminazione diretta, in cui l'algoritmo utilizza esplicitamente una variabile vietata, alla discriminazione indiretta, in

---

<sup>155</sup> GAUDIO, G., *Le discriminazioni algoritmiche*, in *Lavoro Diritti Europa*, n. 1, 2024, p. 3.

<sup>156</sup> *Ibidem*.

<sup>157</sup> *Ivi*, p. 4.

cui criteri solo in apparenza neutrali producono effetti sproporzionatamente negativi su determinate categorie di soggetti. La portata e la rapidità con cui le decisioni algoritmiche vengono applicate amplificano ulteriormente il problema, trasformando errori locali in discriminazioni sistemiche, capaci di incidere su larga scala<sup>158</sup>.

Per queste ragioni, il dibattito contemporaneo non si limita più a interrogarsi sulla rilevanza dei bias algoritmici, ma si concentra sempre più su come prevenirli, mitigarli e correggerli, attraverso una combinazione di strumenti tecnici, giuridici e organizzativi. L'attenzione si rivolge, in particolare, alla trasparenza dei processi decisionali, alla possibilità di audit indipendenti e alla responsabilizzazione (*accountability*) dei soggetti coinvolti nella progettazione, nell'implementazione e nell'utilizzo di tali sistemi<sup>159</sup>.

### **3.2 Come si generano i bias algoritmici**

Il processo di formazione dei bias algoritmici non è mai il prodotto di un singolo errore, ma il risultato di una concatenazione di fattori che si attivano nelle diverse fasi del ciclo di vita di un sistema di intelligenza artificiale. Durante la fase di raccolta dei dati, il rischio maggiore consiste nell'assorbire all'interno dei dataset le distorsioni già presenti nella realtà sociale. I dati di addestramento, infatti, non costituiscono mai una rappresentazione neutra, ma riflettono inevitabilmente le scelte compiute nella loro selezione: quali fonti privilegiare, quali variabili includere, come codificarle e con quale livello di dettaglio<sup>160</sup>. Ogni decisione di questo tipo contribuisce ad orientare il modello, imprimendo un'impronta che può risultare più o meno inclusiva. Se il dataset risulta sbilanciato, incompleto o costruito sulla base di decisioni passate già condizionate da pregiudizi, come nel caso dell'assegnazione di prestiti o dei processi di selezione del personale, tali distorsioni vengono trasferite automaticamente al modello, il quale finisce per riprodurle, e in alcuni casi, per amplificarle. In questo modo, ciò che nasce come un semplice errore di raccolta dati può trasformarsi in una discriminazione sistematica mascherata da apparente oggettività.

---

<sup>158</sup> GAUDIO, G., op. cit., p. 6.

<sup>159</sup> Ivi, p. 8.

<sup>160</sup> PIZZETTI, F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli, Torino, 2021, p. 119.

Nella fase di modellazione, i bias possono emergere da decisioni progettuali che solo in apparenza hanno natura meramente tecnica, ma che in realtà hanno implicazioni sostanziali. La scelta delle variabili esplicative, le modalità con cui vengono attribuiti i pesi, le metriche adottate per misurare la performance o i criteri utilizzati per l'ottimizzazione del modello incorporano sempre assunzioni implicite sugli obiettivi e sulle priorità del sistema<sup>161</sup>. Anche il ricorso a variabili proxy rappresenta un punto particolarmente delicato: si tratta infatti di variabili che, pur non identificando direttamente caratteristiche protette quali etnia, genere o religione, vi risultano strettamente correlate e finiscono per agire come sostituti nascosti, consentendo di fatto all'algoritmo di riprodurre distinzioni formalmente vietate.

Successivamente, nella fase di implementazione (*deployment*), ulteriori distorsioni possono derivare dal disallineamento tra il contesto in cui l'algoritmo è stato addestrato e quello in cui viene effettivamente applicato. Fenomeni come il *concept drift*, ossia l'evoluzione del significato statistico delle variabili nel tempo, e i feedback loop, attraverso i quali le decisioni algoritmiche influenzano i dati futuri alimentando loro stesse, possono contribuire a consolidare e rafforzare schemi discriminatori<sup>162</sup>. Un esempio emblematico è quello dei sistemi di polizia predittiva: addestrati su dati storici già distorti, essi tendono a concentrare l'attenzione sempre sulle medesime aree geografiche, generando nuovi dati che non fanno altro che confermare e amplificare il pregiudizio iniziale.

Il tratto distintivo di questi processi risiede nella loro opacità strutturale: il bias non si manifesta come il risultato di una scelta intenzionale, ma come un effetto emergente dall'interazione tra dati e modello matematico. Proprio questa caratteristica rende più complesso identificarlo e correggerlo, alimentando al contempo l'erronea percezione che la tecnologia si neutrale ed oggettiva<sup>163</sup>.

È proprio da questa opacità che scaturiscono molte delle problematiche specifiche che verranno affrontate nei paragrafi successivi, nei quali saranno analizzati i meccanismi

---

<sup>161</sup>PIZZETTI, F., op. cit., p. 120.

<sup>162</sup>Ivi, p. 121.

<sup>163</sup>Ivi, p. 122.

concreti attraverso cui i bias si producono e si consolidano: dalle disuguaglianze incorporate nei dati storici, alla discriminazione indiretta mediata da variabili proxy, fino al ruolo delle piattaforme digitali e alla loro presunta neutralità.

### 3.3 Dati storici e origine delle distorsioni algoritmiche

Tra le forme più insidiose e pervasive di distorsione algoritmica vi sono le disuguaglianze apprese dai dati storici. Questo fenomeno che si manifesta quando un sistema di *machine learning* viene addestrato su insiemi di dati che non riflettono una realtà neutra, ma incorporano, spesso in maniera silente, pregiudizi, squilibri e stereotipi radicati nella società di riferimento.<sup>164</sup>

Il problema diventa particolarmente evidente in quei settori professionali o sociali caratterizzati da forti asimmetrie di rappresentanza. Laddove, ad esempio, una determinata professione sia stata storicamente dominata da uomini, da individui appartenenti ad una specifica etnia o a un determinato ceto socioeconomico, i dati raccolti rifletteranno inevitabilmente questa sproporzione<sup>165</sup>. Un algoritmo addestrato su tali basi, tenderà a riprodurre, e talvolta ad amplificare, questi squilibri, finendo per privilegiare i profili appartenenti al gruppo maggioritario e per penalizzare, in modo sistematico e invisibile, le minoranze.

Ad aggravare ulteriormente il quadro intervengono i cicli di **retroalimentazione** (*feedback loop*). Quando un algoritmo, già addestrato su dati distorti, viene applicato per prendere decisioni reali, produce nuovi dati che riflettono e rinforzano le stesse distorsioni iniziali. Questi dati, se reintrodotti nel sistema come materiale di addestramento, innescano un effetto cumulativo capace di trasformare pregiudizi locali in vere e proprie discriminazioni su larga scala<sup>166</sup>. L'opacità tecnica, che caratterizza molte architetture di intelligenza artificiale, contribuisce ad alimentare la

---

<sup>164</sup>HEILMAN, M. E., CALEO, S., MANZI, F., *Women at Work: Pathways from Gender Stereotypes to Gender Bias and Discrimination*, in *Annual Review of Organizational Psychology and Organizational Behavior*, vol. 11, 2024, p. 2.

<sup>165</sup>Ivi, p. 4.

<sup>166</sup>Ivi, p. 7.

falsa percezione neutralità tecnologica, ostacolando il controllo e la correzione di tali fenomeni.<sup>167</sup>

Questi meccanismi, apparentemente tecnici e neutrali, hanno dunque un impatto diretto sulla vita delle persone, poiché orientano processi decisionali cruciali come l'accesso al credito, alle cure sanitarie o alle opportunità professionali. È proprio in quest'ultimo ambito che si colloca uno dei casi più emblematici a livello internazionale: l'esperimento condotto da Amazon nel campo del *recruiting*, divenuto un esempio paradigmatico delle derive discriminatorie che possono derivare dall'uso acritico dei dati storici.

### 3.4 Il caso Amazon

Il caso Amazon rappresenta uno degli esempi più citati, a livello internazionale, esempio emblematico di come le disuguaglianze incorporate nei dati storici possano tradursi in forme sistemiche di discriminazione algoritmiche. Nel 2014, l'azienda avviò un progetto sperimentale volto a sviluppare uno strumento di *recruiting automatizzato* basato su tecniche di *machine learning*. L'obiettivo dichiarato era duplice: da un lato velocizzare e standardizzare la valutazione dei curricula, dall'altro ridurre i potenziali pregiudizi soggettivi legati all'intervento umano.<sup>168</sup>

L'architettura del sistema prevedeva l'analisi di decine di migliaia di curricula archiviati da Amazon nell'arco di dieci anni, con lo scopo di attribuire un punteggio di idoneità a ciascun candidato sulla base della somiglianza con i profili che in passato erano stati assunti in ruoli tecnici. In questa logica, i dati pregressi venivano assunti come indicatore affidabile di "successo" professionale in azienda<sup>169</sup>. Tuttavia, tale impostazione trascurava un aspetto decisivo: nel periodo considerato, la maggior parte delle assunzioni

---

<sup>167</sup> HEILMAN, M. E., CALEO, S., MANZI, F., op. cit., p. 8.

<sup>168</sup> TRENTMAN, S., LARWOOD, L., *Gender Discrimination and the Workplace: An Examination of Rational Bias Theory*, in *Sex Roles*, vol. 38, nn. 1-2, 1998, p. 1.

<sup>169</sup> *Ibidem*.

tecniche di Amazon aveva riguardato candidati di sesso maschile, riflettendo la storica sottorappresentanza femminile nel settore tecnologico<sup>170</sup>.

Di conseguenza, l'algoritmo finì così per apprendere, come regola statistica, che il genere maschile costituiva un predittore positivo di idoneità, penalizzando sistematicamente i curricula femminili. Emersero casi paradossali: i modelli abbassavano il punteggio di documenti che contenevano riferimenti espliciti al genere femminile, ad esempio espressioni come "women" in contesti come "women's chess club captain", o che provenivano da istituti scolastici femminili<sup>171</sup>. Si trattava di una forma di discriminazione indiretta, frutto dell'associazione dall'associazione implicita tra certe parole -chiave e gli esiti storici di successo, e non di una scelta intenzionale dei programmatori.<sup>172</sup>

Nonostante i diversi tentativi di rimuovere le variabili esplicitamente discriminatorie (ad esempio, eliminando i termini collegati al genere dal processo di valutazione) il bias continuava a manifestarsi attraverso correlazioni indirette: il modello "ricostruiva" l'informazione di genere combinando altri attributi. Questo riflette una caratteristica strutturale nel *machine learning*: l'eliminazione di una singola variabile non cancella necessariamente l'informazione da essa veicolata, che può essere dedotta da combinazioni di altri attributi.<sup>173</sup>

Un'ulteriore criticità emerse sul piano della valutazione delle performance. I test di accuratezza condotti in fase di sviluppo non prevedevano un'analisi disaggregata per genere, rendendo difficile individuare tempestivamente il problema. Solo con l'applicazione del sistema in contesti reali la distorsione divenne evidente, evidenziando una volta ancora l'assenza di un'adeguata procedura di *bias auditing* in fase di progettazione.<sup>174</sup>

---

<sup>170</sup>TRENTMAN, S., LARWOOD, L., op. cit., p. 2.

<sup>171</sup> Ibidem.

<sup>172</sup>HEILMAN, M. E., CALEO, S., MANZI, F., *Women at Work: Pathways from Gender Stereotypes to Gender Bias and Discrimination*, op. cit., p. 5.

<sup>173</sup> Reuters, *Amazon scraps secret AI recruiting tool that showed bias against women*, 11 ottobre 2018, disponibile su: <https://www.reuters.com>, ultimo accesso: 1 settembre 2025.

<sup>174</sup> Ibidem.

Alla luce della persistenza del fenomeno e all'impossibilità di assicurare un processo di selezione equo, Amazon decise di interrompere definitivamente il progetto. L'episodio dimostra come l'affidamento a dati storici sbilanciati possa compromettere gli obiettivi di neutralità e imparzialità degli algoritmi, trasformando la promessa di efficienza in un meccanismo di esclusione sistemica. Allo stesso tempo, mette in evidenza la necessità di un approccio realmente proattivo alla prevenzione dei bias, fondato su controlli periodici, metriche di equità e strumenti di audit indipendenti da integrare sin dalle prime fasi di sviluppo.<sup>175</sup>

### 3.5 Bias algoritmici nel settore sanitario

L'analisi del caso Amazon dimostra come i sistemi di intelligenza artificiale possano non solo riprodurre, ma persino amplificare le disuguaglianze già presenti nei dati storici, anche in assenza di una volontà discriminatoria. Se effetti di questo tipo si verificano in un ambito come quello del *recruiting*, le loro conseguenze diventano ancor più gravi in settori sensibili come la sanità, dove le decisioni algoritmiche incidono direttamente sull'accesso alle cure e sulla tutela del diritto fondamentale alla salute<sup>176</sup>.

Nel settore sanitario, le distorsioni derivano spesso dall'impiego di dataset clinici che non rappresentano in modo equilibrato la diversità della popolazione. Ad esempio, banche dati costruite prevalentemente su pazienti appartenenti a un determinato gruppo etnico, a uno specifico sesso o ad una particolare fascia d'età tendono a generare modelli meno accurati per soggetti con caratteristiche differenti.<sup>177</sup>

Tali limiti vengono ulteriormente aggravati da fattori come la **sottorappresentazione** di determinati gruppi nei dati di addestramento, che porta a diagnosi o raccomandazioni terapeutiche meno precise, e dalla **disomogeneità qualitativa** delle informazioni

---

<sup>175</sup> Reuters, op. cit.

<sup>176</sup> LONGATO, E., *Bias e fairness nelle applicazioni cliniche dell'intelligenza artificiale*, Tesi di laurea in Ingegneria Biomedica, Dipartimento di Ingegneria dell'Informazione, a.a. 2023–2024, p. 1.

<sup>177</sup> Ibidem.

raccolte, spesso legata a disuguaglianze strutturali nell'accesso alle strutture sanitarie o a differenze nei protocolli di raccolta dei dati.<sup>178</sup>

Un ulteriore elemento di criticità è rappresentato dall'uso di variabili *proxy*, come il codice postale o il reddito. Sebbene queste non menzionano direttamente caratteristiche protette, esse finiscono per agire come indicatori indiretti di fattori socioeconomici e culturali, riproducendo così disparità già radicate nella società.<sup>179</sup>

Particolarmente problematico è l'impiego dell'IA nella predizione del rischio clinico e nella prioritizzazione dei pazienti. Algoritmi destinati ad allocare risorse sanitarie, stabilire liste d'attesa o individuare soggetti ad alto rischio possono sottovalutare la gravità delle condizioni cliniche di determinati gruppi, specialmente se i dati storici riflettono un minore ricorso alle cure da parte loro. In questi casi, il bias non compromette solo l'equità del sistema, ma incide direttamente sulla salute pubblica, con conseguenze potenzialmente gravi in termini di sopravvivenza dei pazienti.<sup>180</sup>

A ciò si aggiunge il nodo dell'opacità: la complessità tecnica dei modelli, combinata alla scarsa trasparenza sulle fonti e sui processi di addestramento, rende estremamente difficile per medici, pazienti e autorità sanitarie comprendere i meccanismi sottesi agli strumenti utilizzati. Ne risulta un enorme ostacolo alla possibilità di contestare decisioni potenzialmente discriminatorie e, soprattutto, un indebolimento della fiducia nell'impiego dell'IA in medicina.<sup>181</sup>

Per queste ragioni, il tema del *bias* in sanità richiede un approccio particolarmente rigoroso. È necessario integrare strategie di *fairness by design*, audit indipendenti, regolamentazione mirata, oltre a favorire il coinvolgimento multidisciplinare di esperti clinici, informatici, giuristi ed eticisti sin dalle prime fasi di sviluppo dei sistemi<sup>182</sup>. Solo un simile approccio consente di ridurre i rischi, garantire equità ed evitare che l'innovazione tecnologica si trasformi in un ulteriore fattore di esclusione.

---

<sup>178</sup>LONGATO, E., op. cit., p. 1.

<sup>179</sup>Ivi, p. 2.

<sup>180</sup>Ivi, p. 2-3.

<sup>181</sup>Ivi, p. 3.

<sup>182</sup>Ivi, p. 4.

### 3.6 Verso soluzioni efficaci: prevenzione, controllo e responsabilità nelle discriminazioni algoritmiche

La complessità delle discriminazioni algoritmiche in ambito lavorativo impone un ripensamento profondo degli strumenti giuridici e organizzativi attualmente disponibili<sup>183</sup>. La progressiva diffusione di sistemi di *algorithmic management*, spesso opachi e difficilmente comprensibili persino agli stessi operatori, ha reso evidente il rischio di una riproduzione amplificata dei pregiudizi presenti nei dati storici<sup>184</sup>. Non si tratta soltanto di un problema tecnico: è una questione che incide direttamente sui diritti fondamentali dei lavoratori e sulla tenuta complessiva delle garanzie poste a presidio dell'uguaglianza.<sup>185</sup>

In questo contesto, la prevenzione assume un ruolo centrale. Lo strumento principale è *Data Protection Impact Assessment*, previsto dall'art. 35 GDPR.<sup>186</sup> Si tratta di una valutazione d'impatto obbligatoria nei casi in cui il trattamento dei dati presenti rischi elevati per i diritti e le libertà degli interessati. Attraverso il DPIA, il titolare del trattamento è chiamato a mappare i processi, individuare i possibili rischi, inclusi quelli di discriminazione algoritmica, e predisporre misure concrete per ridurli. Non basta, quindi una compilazione meramente formale: il DPIA deve essere uno strumento sostanziale di prevenzione, capace di anticipare criticità prima che queste si traducano in effetti discriminatori nei confronti dei lavoratori. Proprio per questo, la valutazione deve essere aggiornata periodicamente, soprattutto quando riguarda sistemi di *machine learning* che, modificando nel tempo i propri criteri decisionali, possono generare nuovi rischi non previsti in origine.<sup>187</sup>

---

<sup>183</sup>GAUDIO, G., *Le discriminazioni algoritmiche*, in *Numero 1/2024*, Università degli Studi di Torino, p. 90.

<sup>184</sup> Ivi, p. 93

<sup>185</sup> Ibidem.

<sup>186</sup> Art. 35, Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali nonché alla libera circolazione di tali dati (General Data Protection Regulation – GDPR)

<sup>187</sup>GAUDIO, G., op. cit., p. 94.

Accanto al DPIA, un altro presidio essenziale è costituito dai diritti di informazione e di accesso: essi consentono al lavoratore e alle sue rappresentanze di conoscere le logiche di funzionamento dell'algoritmo e la tipologia di dati trattati, favorendo così un controllo effettivo da parte dei lavoratori e delle loro rappresentanze<sup>188</sup>. La trasparenza non è solo un principio astratto, ma una condizione per esercitare un controllo effettivo e consapevole sulle decisioni automatizzate.

Infine, la verifica esterna è garantita dagli *audit* indipendenti e dai meccanismi di certificazione. Questi strumenti hanno il compito di monitorare costantemente le qualità dei dati, di accertare l'assenza di *bias* e di verificare la conformità alle norme vigenti, riducendo il rischio che i sistemi di *algorithmic management* operino in maniera opaca o distorsiva.<sup>189</sup>

La repressione delle discriminazioni algoritmiche, per essere davvero efficace, deve poggiarsi su strumenti capaci di riequilibrare l'asimmetria informativa che separa datore di lavoro e lavoratore. In questo senso, l'inversione o l'attenuazione dell'onere della prova, già prevista nelle direttive antidiscriminatorie e riconosciuta dalla giurisprudenza europea, assume un ruolo decisivo. Essa consente infatti di valorizzare come elemento probatorio la mancata collaborazione della parte datoriale<sup>190</sup>. Non a caso, La Corte di Giustizia, nelle pronunce *Meister*<sup>191</sup> e *Danfoss*<sup>192</sup>, ha chiarito che l'assenza di informazioni o la mancata trasparenza possono costituire indizi gravi dell'esistenza di una discriminazione<sup>193</sup>. A ciò si aggiunge la possibilità per il giudice, di ordinare l'esibizione di documenti tecnici, come il DPIA, e disporre consulenze tecniche *percipientes*, così da ricostruire il funzionamento effettivo del sistema algoritmico.<sup>194</sup>

Un ulteriore ambito di intervento particolarmente rilevante è quello della contrattazione collettiva. Con questa espressione si indica l'insieme degli accordi stipulati tra

---

<sup>188</sup> GAUDIO, G., op. cit., p. 112.

<sup>189</sup> Ivi, p. 9.

<sup>190</sup><sup>190</sup>Ivi, p.125.

<sup>191</sup> CGUE, causa C-415/10, *Meister* [2012], ECLI:EU:C:2012:217.

<sup>192</sup> CGCE, causa C-109/88, *Handels- og Kontorfunktionærernes Forbund i Danmark (Danfoss)* [1989], ECLI:EU:C:1989:383.

<sup>193</sup> Ivi, pp. 134-135.

<sup>194</sup> Ivi, p. 140.

organizzazioni sindacali e datori di lavoro, finalizzati a disciplinare in maniera uniforme i rapporti di lavoro in un determinato settore o ambito aziendale. Essa rappresenta uno strumento fondamentale per integrare e rafforzare le tutele legali, poiché consente di introdurre regole condivise e vincolanti che vanno oltre la legislazione generale, adattandole alle specificità del contesto produttivo. Alcune esperienze giurisprudenziali recenti, come quelle dei Tribunali di Bologna e Palermo, hanno dimostrato come l'azione sindacale possa rivelarsi decisiva per portare alla luce e correggere pratiche discriminatorie generate dall'uso di algoritmi<sup>195</sup>. La dimensione collettiva, infatti, offre la possibilità di affrontare i problemi in modo sistematico, superando i limiti inevitabili delle sole iniziative individuali e garantendo una tutela più ampia ed efficace.<sup>196</sup>

In prospettiva, *de iure condito*, diventa imprescindibile garantire un'effettiva applicazione delle DPIA, in modo da renderli strumenti concreti di prevenzione e non meri adempimenti formali. A ciò si devono affiancare la tracciabilità delle decisioni, un accesso realmente sostanziale alle informazioni e meccanismi di verifica tecnica costanti, capaci di assicurare un controllo continuativo sul funzionamento degli algoritmi.<sup>197</sup>

Sul piano *de iure condendo*, sarebbe opportuno invece rafforzare il sistema delle tutele. In particolare, potrebbe essere conveniente estendere l'obbligo di valutazione preventiva a tutti i sistemi di *algorithmic management* ad alto impatto, introdurre l'obbligatorietà degli audit indipendenti e sancire in maniera chiara, anche a livello normativo, criteri precisi per l'inversione dell'onere della prova nei casi in cui l'opacità tecnica renda difficile accertare il funzionamento del processo decisionale.<sup>198</sup>

L'algoritmo, in conclusione, non è intrinsecamente discriminatorio: la sua equità dipende dal contesto regolativo, dalle scelte progettuali e dall'efficacia dei controlli cui è sottoposto.<sup>199</sup> Se governato secondo principi di trasparenza, correttezza e responsabilità,

---

<sup>195</sup>GAUDIO, G., *op. cit.*, p.149.

<sup>196</sup>Ibidem

<sup>197</sup> Ibidem

<sup>198</sup> Ivi, p. 152.

<sup>199</sup>Ivi, p. 150.

esso può divenire uno strumento capace di ridurre pregiudizi e disuguaglianze, favorendo un'organizzazione del lavoro più equa e inclusiva.

Conclusivamente, la sfida non consiste nel frenare l'innovazione tecnologica, ma nell'orientarla verso modelli capaci di coniugare efficienza ed effettiva tutela dei diritti fondamentali. Solo in questo modo l'uso degli algoritmi potrà trasformarsi da fattore di rischio a leva di progresso, garantendo che la tecnologia sia posta al servizio della dignità e dell'uguaglianza dei lavoratori.

Se la lotta ai *bias* costituisce un primo fronte di tutela, il fenomeno più ampio della governamentalità algoritmica rivela come l'intelligenza artificiale non si limiti a influenzare singole decisioni, ma riorganizzi in profondità i rapporti di potere all'interno del lavoro contemporaneo.

#### **4. Tecnologie algoritmiche e trasformazione del rapporto di lavoro**

La progressiva penetrazione dell'intelligenza artificiale nei processi produttivi e organizzativi non può essere compresa unicamente come un fenomeno di innovazione tecnologica<sup>200</sup>. Essa si inserisce in una trasformazione più profonda delle modalità di esercizio del potere nella società contemporanea, che segna il passaggio dalla sorveglianza tradizionale a quella che alcuni autori, come Antoinette Rouvroy e Bernard Stiegler, definiscono governamentalità algoritmica.<sup>201</sup>

In questo nuovo paradigma, la raccolta pervasiva di dati e il loro trattamento tramite sistemi predittivi consentono non soltanto di monitorare il comportamento umano, ma arrivano a modularlo e orientarlo in tempo reale, incidendo sulle scelte individuali in modo impercettibile e costante.<sup>202</sup>

---

<sup>200</sup>ZAPPALÀ, L., *Informatizzazione dei processi decisionali e diritto del lavoro: algoritmi, poteri datoriali e responsabilità del prestatore nell'era dell'intelligenza artificiale*, in *Biblioteca "20 Maggio"*, n. 2, 2021, p. 1.

<sup>201</sup>PIETROPAOLI, S., *Dalla sorveglianza al controllo: la parabola della governamentalità algoritmica*, in *Rivista Italiana di Informatica e Diritto*, n. 2, 2024, p. 25.

<sup>202</sup> Ivi, p. 27.

Applicata al lavoro questa dinamica determina una ridefinizione radicale dei rapporti tra datatore di lavoro, lavoratore e piattaforma tecnologica si ridefiniscono radicalmente<sup>203</sup>: L'attività produttiva viene progressivamente "datificata" ossia tradotta in unità misurabili, analizzabili e ottimizzabili.<sup>204</sup> La prestazione lavorativa, a sua volta, diventa oggetto di valutazioni automatizzate, profilazioni e suggerimenti comportamentali.<sup>205</sup>

Questa transizione porta con sé non solo un incremento dell'efficienza e della personalizzazione, ma anche l'emersione di nuove forme di controllo sociale, spesso opache e difficili da contestare.<sup>206</sup> Ne derivano asimmetrie di potere sempre più marcate<sup>207</sup>, una riduzione dell'autonomia professionale<sup>208</sup> e il rischio concreto di discriminazione algoritmiche.<sup>209</sup>

Comprendere il fenomeno della governamentalità algoritmica significa interrogarsi sulle implicazioni di un potere diffuso e invisibile, che non si esercita più attraverso ordini espliciti o sanzioni dirette, ma mediante la previsione e la modulazione delle condotte<sup>210</sup>. In questo nuovo scenario, la sfida non è soltanto normativa, ma soprattutto culturale: occorre sviluppare nuovi strumenti concettuali, competenze digitali e forme di consapevolezza<sup>211</sup> che permettano ai lavoratori di comprendere e governare, piuttosto che subire, le logiche algoritmiche che modellano il loro operato e le loro prospettive professionali.<sup>212</sup>

Questa trasformazione si colloca all'interno della più ampia cornice della quarta rivoluzione industriale, segnata dall'integrazione pervasiva di tecnologie digitali avanzate.<sup>213</sup> Essa ha impresso un'accelerazione senza precedenti ai processi produttivi e organizzativi, incidendo profondamente sulla natura e sulle modalità di svolgimento del

---

<sup>203</sup>ZAPPALA', L., op. cit., p.3

<sup>204</sup>Ivi, p. 4.

<sup>205</sup>Ivi, pp. 7-8

<sup>206</sup>PIETROPAOLI, S., op. cit., p. 29.

<sup>207</sup>Loredana Zappalà, op. cit., p. 8.

<sup>208</sup>Ivi, p. 9.

<sup>209</sup>Ibidem

<sup>210</sup>PIETROPAOLI, S., op. cit., pp. 30-31.

<sup>211</sup>ZAPPALA', L., op. cit., p. 9.

<sup>212</sup>Ivi, p. 9.

<sup>213</sup>Ivi, p. 1.

lavoro<sup>214</sup>. L'elemento caratterizzante di questa fase storica è la datificazione sistematica di ogni attività umana<sup>215</sup>, vale a dire la traduzione di azioni, comportamenti e scelte in dati analizzabili e processabili da sistemi algoritmici.

In questo contesto, prende forma l'“algocrazia” ossia il governo dei processi decisionali attraverso algoritmi. Essa si configura come nuova forma di potere gestionale<sup>216</sup>, capace di orientare in maniera automatizzata l'intero ciclo di vita del rapporto di lavoro, dall'assunzione alla cessazione.<sup>217</sup>

Questa transizione segna un vero e proprio mutamento di paradigma rispetto alle strutture tradizionali del diritto del lavoro<sup>218</sup>. Se in passato il potere direttivo e organizzativo del datore si esercitava principalmente attraverso interazioni umane dirette, oggi esso può essere veicolato e amplificato da strumenti tecnologici dotati di capacità predittive e prescrittive.<sup>219</sup> Grazie alla combinazione di *big data*, *machine learning* e intelligenza artificiale,<sup>220</sup> tali sistemi non si limitano più a registrare e monitorare l'attività, ma sono in grado di elaborare modelli decisionali che si sostituiscono, almeno in parte alla valutazione umana, con un livello di autonomia e pervasività tale da mettere in discussione le tradizionali categorie giuridiche di imputazione e responsabilità<sup>221</sup>.

L'impatto di questa trasformazione si riflette in almeno tre ambiti fondamentali.

1. **Ridefinizione del concetto di controllo datoriale:** la supervisione non si limita più alla verifica ex post della prestazione, ma assume la forma di un monitoraggio continuo e in tempo reale<sup>222</sup>, capace di orientare le azioni del lavoratore nel momento stesso in cui vengono compiute<sup>223</sup>.

---

<sup>214</sup>ZAPPALA', L., op. cit., pp. 1-2.

<sup>215</sup> Ivi, p. 4.

<sup>216</sup> Ivi, p. 7.

<sup>217</sup> Ibidem.

<sup>218</sup> Ivi, pp. 7-8.

<sup>219</sup> Ivi, p. 8.

<sup>220</sup> Ibidem.

<sup>221</sup> Ivi, p. 9.

<sup>222</sup> Ivi, p. 8.

<sup>223</sup> Ibidem.

2. **Automazione delle decisioni gestionali:** le scelte relative a valutazione delle performance, all'attribuzione delle mansioni e alla modulazione dei turni<sup>224</sup>, sono sempre più affidate ad algoritmi. Questo produce effetti immediati sulla libertà professionale, sull'autonomia negoziale e sulla possibilità per il lavoratore di incidere sulle proprie condizioni di impiego <sup>225</sup>.
3. **Nuove asimmetrie informative:**<sup>226</sup> il lavoratore non solo si confronta con decisioni opache e difficili da contestare, ma subisce anche l'influenza dei *bias* insiti nei dati su cui l'algoritmo è stato addestrato. In questo modo, la mancanza di trasparenza e la dipendenza dal dato rischiano di accentuare le disuguaglianze già presenti nel mercato del lavoro. <sup>227</sup>

La sfida che oggi si pone il diritto del lavoro è quella di assicurare che le opportunità, offerte dalle nuove tecnologie, in termini di efficienza, sicurezza e personalizzazione della prestazione, non si trasformino in strumenti di compressione dei diritti fondamentali<sup>228</sup>. Per raggiungere tale equilibrio è necessario adottare un approccio regolativo multilivello<sup>229</sup>, capace di intrecciare norme generali e astratte con clausole specifiche dedicate alla gestione algoritmica, ponendo sempre al centro la tutela della dignità e l'autonomia del lavoratore<sup>230</sup>.

Dopo aver analizzato le implicazioni generali della **governamentalità algoritmica** nel contesto occupazionale, è ora possibile soffermarsi sulle sue declinazioni più concrete all'interno di modelli organizzativi e gestionali. Nei paragrafi che seguono, saranno esaminati, in particolare, il fenomeno del *Digital Taylorism*, la ridefinizione del ruolo manageriale, l'automazione decisionale nella valutazione delle performance, nonché le ricadute di tali processi sulla dignità, l'autonomia e la creatività del lavoratore<sup>231</sup>. L'obiettivo è offrire una visione organica delle trasformazioni in atto, capace di integrare

---

<sup>224</sup>ZAPPALA', L., op. cit., pp. 8-9.

<sup>225</sup> Ivi, p. 9.

<sup>226</sup> Ibidem.

<sup>227</sup> Ibidem.

<sup>228</sup> Ibidem.

<sup>229</sup> Ibidem.

<sup>230</sup> Ibidem.

<sup>231</sup> Ivi, p. 4.

la prospettiva giuridica con quella ingegneristica, così da cogliere sia le dinamiche di potere che le dimensioni tecniche dei sistemi di controllo automatizzato.

#### **4.1 Digital Taylorism e ridefinizione del ruolo manageriale**

Il concetto di *Digital Taylorism* rappresenta la trasposizione, in chiave tecnologica, dei principi di organizzazione scientifica del lavoro elaborati da Frederick W. Taylor agli inizi del XX secolo<sup>232</sup>. Se, nel taylorismo classico, l'ottimizzazione della produttività si basava sulla misurazione dei tempi e dei metodi, oggi medesima logica viene applicata mediante strumenti algoritmici, piattaforme digitali e sistemi di monitoraggio costante<sup>233</sup>.

In tale scenario, la prestazione lavorativa viene scomposta in micro-attività facilmente tracciabili, valutabili e comparabili attraverso la raccolta e all'analisi, in tempo reale, di grandi quantità di dati<sup>234</sup>. La progressiva "datificazione" del lavoro non si limita a registrare le attività svolte, ma permette di prevedere i comportamenti futuri del lavoratore, di suggerire azioni correttive e, in alcuni casi, di imporre modifiche operative immediate<sup>235</sup>. In tal modo, l'algoritmo assume un ruolo assimilabile a quello di un supervisore digitale, in grado di intervenire direttamente sui processi produttivi<sup>236</sup>.

L'automazione delle funzioni di supervisione e coordinamento, tradizionalmente affidate al management intermedio, determina due effetti di grande rilievo. Da un lato, riduce la discrezionalità umana nella valutazione delle performance, delegando questo compito ai sistemi automatici<sup>237</sup>. Dall'altro, riconfigura il ruolo del management, che si orienta sempre meno verso la gestione diretta delle risorse umane e sempre più verso il controllo e la regolazione dell'infrastruttura algoritmica<sup>238</sup>. Il manager si trasforma in un mediatore tra la logica impersonale del sistema digitale e le esigenze concrete dei lavoratori, con il

---

<sup>232</sup> ZAPPALA', L., op. cit., p. 4.

<sup>233</sup> Ibidem.

<sup>234</sup> Ivi, p. 5.

<sup>235</sup> Ibidem.

<sup>236</sup> Ivi, p. 6.

<sup>237</sup> PICCINI, M., CARPIGNANO, A., *I fattori umani nel progetto di sistemi di controllo e di interfacce uomo-macchina in sistemi complessi altamente automatizzati*, Politecnico di Torino – Dipartimento di Energetica, p. 2.

<sup>238</sup> Ivi, p. 3.

rischio, tuttavia, di perdere la capacità di cogliere la qualità e la specificità dei contesti operativi.<sup>239</sup>

Questa standardizzazione digitale produce conseguenze rilevanti sull'autonomia professionale e sulla creatività dei lavoratori: la rigidità dei parametri predefiniti riduce gli spazi di iniziativa individuale, mentre l'adeguamento sulle metriche algoritmiche genera una pressione continua e un clima di controllo pervasivo<sup>240</sup>. Sul piano giuridico, il *Digital Taylorism* solleva interrogativi significativi circa compatibilità con i principi fondamentali del diritto del lavoro, in particolare per quanto riguarda la tutela della dignità, della salute psico-fisica e della libertà professionale del lavoratore<sup>241</sup>. La questione non si limita soltanto alla sostituzione delle decisioni umane con quelle algoritmiche, ma investe anche la legittimità di attribuire a sistemi automatizzati funzioni di valutazione, coordinamento e sanzione, che incidono direttamente sulla posizione contrattuale e sulla carriera dei dipendenti<sup>242</sup>.

In definitiva, l'automazione decisionale applicata alla valutazione della performance pone il diritto del lavoro di fronte a una duplice sfida. Da un lato, garantire che gli strumenti algoritmici non degenerino in meccanismi opachi e incontestabili, in grado di incidere unilateralmente sul destino professionale del lavoratore. Dall'altro, assicurare che le garanzie previste dal GDPR, e in particolare il diritto all'intervento umano, trovino una concreta attuazione nella prassi organizzativa.

Solo così è possibile evitare che il controllo digitale si traduca in una forma di neo-taylorismo radicale, capace di comprimere dignità, autonomia e libertà professionale. La riflessione che segue si concentrerà proprio su questi profili critici, analizzando come il ricorso a processi decisionali automatizzati ridisegni non soltanto le dinamiche contrattuali, ma anche le relazioni di potere e i diritti fondamentali coinvolti.

---

<sup>239</sup>PICCINI, M, CARPIGNANO, A., op. cit., p.3

<sup>240</sup>ZAPPALÀ, L., op. cit., p. 6.

<sup>241</sup>PICCINI, M, CARPIGNANO, A., op. cit., p. 4.

<sup>242</sup>ZAPPALÀ, L., op. cit., p. 7.

## 4.2 Automazione decisionale e valutazione della performance

L'introduzione di sistemi algoritmici per la valutazione della performance lavorativa non rappresenta un semplice aggiornamento tecnologico, ma un vero e proprio cambio di rotta nella gestione del rapporto di lavoro<sup>243</sup>. Attraverso l'elaborazione di dati provenienti da diverse fonti: tempi di esecuzione, livelli di produttività, errori commessi, interazioni con clienti, utilizzo di strumenti aziendali, l'algoritmo elabora punteggi e classifiche costantemente aggiornati<sup>244</sup>. Questo monitoraggio continuo, reso possibile dall'interconnessione tra piattaforme digitali e sensori, sostituisce le tradizionali valutazioni periodiche<sup>245</sup>.

Gli output generati non si limitano a fornire un quadro descrittivo dell'attività, ma incidono direttamente su decisioni rilevanti: promozioni, attribuzione di premi di produttività, definizione di percorsi di carriera o, al contrario, attivazione di procedure correttive e disciplinari<sup>246</sup>. Tali esiti, quando derivano da processi completamente automatizzati, pongono un problema di compatibilità con il diritto alla protezione dei dati personali e, in particolare, con l'art. 22 del Regolamento (UE) 2016/679 (GDPR)<sup>247</sup>.

Il **comma 1 dell'art. 22** stabilisce che l'interessato *“ha il diritto di non essere sottoposto a una decisione basata unicamente su un trattamento automatizzato, inclusa la profilazione”* quando tale decisione produce effetti giuridici o incide in modo analogo significativamente sulla persona<sup>6</sup>. Nel contesto lavorativo, per “effetti giuridici” si intende qualsiasi conseguenza che modifichi lo status contrattuale o le condizioni di impiego del lavoratore, mentre l’“incidenza significativa” si concretizza in valutazioni negative che possono pregiudicare opportunità di avanzamento, accesso a bonus o mantenimento del posto di lavoro.<sup>248</sup>

---

<sup>243</sup>ZAPPALA', L., op. cit., p. 7.

<sup>244</sup> Ibidem.

<sup>245</sup> Ivi, p. 8.

<sup>246</sup> Ibidem.

<sup>247</sup> Art. 22, Regolamento (UE) 2016/679 (General Data Protection Regulation – GDPR)

<sup>248</sup> Ivi, comma 1.

Il **comma 2** individua tre eccezioni:

- a) la necessità della decisione per la conclusione o l'esecuzione di un contratto tra lavoratore e datore di lavoro;
- b) l'autorizzazione prevista dal diritto dell'Unione o dello Stato membro, con adeguate misure di salvaguardia;
- c) il consenso esplicito dell'interessato<sup>249</sup>.

Tuttavia, anche in presenza di una di queste deroghe, il **comma 3** impone che il datore di lavoro adotti misure idonee a tutelare i diritti, le libertà e i legittimi interessi del lavoratore. Tali misure includono il diritto all'intervento umano, alla possibilità di esprimere la propria opinione e a contestare la decisione<sup>250</sup>. La giurisprudenza e la dottrina evidenziano che l'"intervento umano" non può essere puramente formale: deve trattarsi di un riesame effettivo, svolto da una persona competente e dotata di reale potere decisionale<sup>251</sup>.

Il contributo di Ilaria Del Giglio ha evidenziato come un sistema di monitoraggio algoritmico, per poter dirsi conforme al quadro normativo e rispettoso dei diritti dei lavoratori, non possa ridursi a un insieme di metriche rigide e standardizzate, ma debba includere margini di valutazione qualitativa e contestuale<sup>11</sup>. In mancanza di tali correttivi, il rischio è quello di consolidare bias già presenti nei dataset storici o, al contrario, di penalizzare condotte virtuose che sfuggono alle categorie predefinite. Tale prospettiva richiama direttamente l'art. 22 GDPR, il quale impone non solo il diritto a non essere sottoposti a decisioni unicamente automatizzate, ma anche l'obbligo per il datore di lavoro di garantire interventi umani effettivi, capaci di riesaminare criticamente le valutazioni algoritmiche e di assumerne la responsabilità giuridica.<sup>252</sup>

L'analisi di Ilaria Del Giglio, incentrata sui rischi di rigidità e cristallizzazione delle metriche, trova un naturale proseguimento nel dibattito sulla trasparenza e l'opacità

---

<sup>249</sup> Regolamento (UE) 2016/679 (GDPR), art. 22, comma 2.

<sup>250</sup> Ivi, comma 3.

<sup>251</sup> DEL GIGLIO, I., *Valutazione della performance mediante tecniche di People Analytics. Privacy in employment, controllo o innovazione*, in *Journal of Ethics and Legal Technologies*, vol. 3, n. 2, novembre 2021, p. 6.

<sup>252</sup> Ivi, p. 4.

algoritmica, così come delineato da Maria Romana Allegri. Se, infatti, la valutazione della performance basata su strumenti di *people analytics* rischia di ridursi a un meccanismo cieco e autoreferenziale, il problema non è solo giuridico ma anche epistemologico: il lavoratore si trova di fronte a decisioni che incidono direttamente sulla sua carriera senza poter accedere a criteri realmente comprensibili. La trasparenza promessa dai sistemi digitali si rivela spesso illusoria, poiché la complessità tecnica e la natura predittiva degli algoritmi rendono arduo, se non impossibile, comprendere il percorso logico che conduce dal dato grezzo alla decisione finale.<sup>253</sup>

Allegri evidenzia come il concetto stesso di trasparenza algoritmica sia ambiguo e contraddittorio. Da un lato, esso viene esaltato come garanzia di accountability e strumento di tutela dei diritti fondamentali; dall'altro, si scontra con la realtà di un linguaggio tecnico opaco, spesso inaccessibile ai non addetti ai lavori. La trasparenza, lungi dall'essere un concetto neutro, assume così una valenza politica e retorica, rischiando di trasformarsi in una "parola d'ordine" che maschera la persistente asimmetria informativa tra chi sviluppa e controlla i sistemi e chi invece li subisce<sup>254</sup>. Nel contesto della valutazione delle performance lavorative, ciò significa che i punteggi e le classifiche prodotti dagli algoritmi, pur presentati come oggettivi e imparziali, restano di fatto opachi per i lavoratori, che non dispongono degli strumenti per comprenderne né la logica né gli eventuali errori di categorizzazione.

Un ulteriore nodo critico riguarda il rapporto tra trasparenza e linguaggio. Allegri sottolinea come la traduzione di logiche algoritmiche complesse in forme comunicabili al pubblico finisca per produrre inevitabili semplificazioni, che spesso svuotano il concetto stesso di trasparenza<sup>255</sup>. Nella pratica, le informative fornite ai lavoratori, anche quando formalmente conformi al GDPR, rischiano di ridursi a enunciazioni astratte, prive di reale capacità esplicativa. Ciò genera una trasparenza solo apparente, che non consente né al singolo lavoratore di difendersi da valutazioni ingiuste, né alle istituzioni di esercitare un controllo effettivo. Il divario tra la promessa di chiarezza e la persistenza

---

<sup>253</sup> ALLEGRI, M.R., *op. cit.*, p. 3.

<sup>254</sup> Ivi, p. 5.

<sup>255</sup> Ivi, pp. 7-8.

dell'opacità mina così l'effettività delle tutele previste dall'art. 22 GDPR, già messe in discussione dalla crescente automazione delle decisioni.

In questa prospettiva, la riflessione di Allegri porta a riconsiderare criticamente il principio stesso di trasparenza come strumento regolativo. Non basta, infatti, invocare maggiore chiarezza sugli algoritmi: occorre interrogarsi su quale tipo di trasparenza sia realmente utile e praticabile. L'autrice suggerisce che la soluzione non risiede in una illusoria totale accessibilità del codice o dei modelli, ma piuttosto nello sviluppo di pratiche di controllo e verifica sociale, in grado di tradurre la complessità tecnica in criteri di responsabilità comprensibili e giuridicamente rilevanti<sup>256</sup>. Solo attraverso questo approccio è possibile evitare che la valutazione algoritmica della performance lavorativa si traduca in una forma di "opacità legittimata", capace di erodere la dignità e l'autonomia dei lavoratori pur mantenendo un'apparenza di legalità.

### **4.3 Impatto sull'autonomia e sulla dignità del lavoratore**

Come osserva Covelli, la riflessione sull'autonomia decisionale e sulla dignità del lavoratore non può essere disgiunta da quella sulla trasparenza e sulla conoscibilità algoritmica, intese come vere garanzie di libertà. La storia dello sviluppo informatico, oscillante tra logiche di apertura e di chiusura, tra cooperazione e segreto industriale, mostra infatti come l'ideale della trasparenza abbia spesso lasciato spazio a modelli proprietari e opachi, nei quali il potere di controllo è concentrato nelle mani di pochi attori economici.<sup>257</sup>

Questo passaggio dall'apertura alla chiusura, dal libero accesso alla protezione tramite segreto e proprietà intellettuale, non è soltanto una dinamica tecnica o commerciale, ma rappresenta un nodo giuridico cruciale: anche nel rapporto di lavoro, i sistemi digitali nascono con una promessa di trasparenza, ma finiscono per operare in modo strutturalmente opaco, condizionando autonomia e diritti dei prestatori.

---

<sup>256</sup> ALLEGRI, M. R., *L'ambiguo principio (anche costituzionale?) della trasparenza algoritmica fra tecnologia, diritto e linguaggio*, in *Saggi*, 2021, p. 12.

<sup>257</sup> COVELLI, L., *Lavoro e intelligenza artificiale: dalla trasparenza alla conoscibilità*, in *LLI*, vol. 9, n. 1, 2023, pp. 79–80.

In questo quadro, la privacy non può più essere intesa nella sua concezione originaria di diritto “negativo”, volto esclusivamente a difendere una sfera individuale da intrusioni esterne. Come sottolinea Covelli, nel contesto digitale essa si declina sempre più come autodeterminazione informativa e come *habeas data*, cioè diritto della persona a mantenere il controllo effettivo sui propri dati<sup>258</sup>. Questa trasformazione concettuale ha trovato riconoscimento nella Carta dei diritti fondamentali dell’Unione europea, che distingue tra diritto al rispetto della vita privata (art. 7) e diritto alla protezione dei dati personali (art. 8), attribuendo a quest’ultimo una dimensione dinamica, capace di seguire la circolazione dei dati e di imporre precisi obblighi a chi li tratta.<sup>259</sup>

La centralità del dato personale, divenuto risorsa economica e base per decisioni automatizzate, porta con sé il rischio di nuove forme di stigmatizzazione ed esclusione sociale. L’automazione dei processi decisionali, tipica dell’IA applicata al contesto lavorativo, può infatti determinare valutazioni che incidono sulla reputazione professionale e sulla stessa possibilità di accesso e permanenza nel lavoro. Di fronte a questo scenario, il consenso al trattamento dei dati, base giuridica tradizionalmente prevista dal GDPR, non può essere considerato sufficiente: se reso in un contesto di forte asimmetria di potere, come quello lavorativo, esso rischia di trasformarsi in una formalità, incapace di garantire una protezione sostanziale<sup>260</sup>.

Questi elementi dimostrano come la protezione della dignità e dell’autonomia del lavoratore debba passare necessariamente attraverso la governance dei dati personali, che si pone come strumento di bilanciamento tra esigenze organizzative e tutela dei diritti fondamentali. È proprio in questa direzione che si muove il paragrafo successivo, dedicato a esaminare la disciplina della privacy e della sorveglianza algoritmica nel contesto lavorativo, con particolare attenzione ai principi del GDPR, al trattamento dei dati biometrici e comportamentali, e agli obblighi informativi che gravano sul datore di lavoro.

---

<sup>258</sup> COVELLI, L., op. cit., p. 81.

<sup>259</sup> Ibidem.

<sup>260</sup> Ivi, pp. 81–82.

#### 4.4 Privacy e sorveglianza algoritmica nel contesto lavorativo

L'impiego dell'intelligenza artificiale nei processi di gestione del personale, con particolare riguardo alle procedure di reclutamento e valutazione, ha introdotto modalità di trattamento dei dati personali sempre più pervasive. Tali pratiche includono non solo dati anagrafici e professionali, ma anche informazioni di natura biometrica, sanitaria e comportamentale, con conseguenze rilevanti sulla tutela della sfera privata del lavoratore.

Come sottolinea Sarra, la biometria, definita dall'art. 4, n. 14, GDPR<sup>261</sup> come il trattamento di dati personali relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona, che ne permettono o confermano l'identificazione univoca, può essere distinta in due categorie. Da un lato vi è la biometria **“hard”**, basata su tratti fisici stabili e immutabili nel tempo come le impronte digitali, la geometria facciale, il riconoscimento dell'iride o il profilo genetico del DNA. Dall'altro si colloca la **biometria “soft”**, che invece fonda su parametri più dinamici e mutevoli, legati al comportamento ed al contesto come la postura, il timbro e le inflessioni della voce, le espressioni facciali, la modalità di digitazione o l'andatura<sup>262</sup>.

Nel contesto lavorativo, la biometria “soft” si presta ad un utilizzo particolarmente invasivo, in quanto consente ai sistemi di intelligenza artificiale di estrarre, attraverso analisi predittive, indicatori di personalità, stati emotivi o attitudini comportamentali. L'obiettivo dichiarato è spesso quello di costruire un “profilo ideale” del lavoratore o del candidato. Tuttavia, pratiche di questo tipo espongono a rischi concreti di discriminazioni indirette, specie quando i criteri utilizzati derivano da dati storici distorti o culturalmente parziali<sup>263</sup>. Inoltre, la natura stessa delle informazioni trattate, che possono riflettere aspetti sanitari o comportamentali o intimi della sfera comportamentale, incide profondamente sulla vita privata, ponendo l'accento sulla necessità di rispettare vincoli stringenti di liceità o proporzionalità.

---

<sup>261</sup> Regolamento (UE) 2016/679, art. 4, n. 14

<sup>262</sup> SARRA, C., *L'uso di dati biometrici nelle procedure di reclutamento al lavoro mediante strumenti di Intelligenza Artificiale. Difficoltà normative multilivello*, in *Journal of Ethics and Legal Technologies*, vol. 4, n. 2, novembre 2022, p. 49

<sup>263</sup>Ivi, pp. 50-51.

L'art 5 GDPR sancisce i principi fondamentali che devono orientare qualsiasi trattamento dei dati personali: liceità, correttezza e trasparenza, limitazione delle finalità; minimizzazione dei dati; esattezza; limitazione alla conservazione; integrità e riservatezza.<sup>264</sup>In ambito lavorativo, tali principi assumono maggiore importanza, poiché l'asimmetria di potere tra datore di lavoro e prestatore può rendere molto fragile il consenso, fino a svuotarlo di reale libertà di scelta. Non a caso, la dottrina e le autorità di controllo hanno spesso sottolineato come il consenso, in questo specifico contesto, rappresenti un fondamento giuridico debole o addirittura inidoneo.

Per quanto riguarda i dati particolari, tra cui rientrano i dati biometrici e sanitari, l'art 9 GDPR pone un divieto generale di trattamento, a tutela della sfera più intima della persona, salvo che ricorrano specifiche deroghe. Il legislatore europeo infatti ammette alcune eccezioni. In particolare, il trattamento è consentito quando:

- l'interessato ha prestato un **consenso esplicito** per una o più finalità determinate;
- il trattamento è necessario per **assolvere obblighi ed esercitare diritti** del titolare o dell'interessato in materia di diritto del lavoro, sicurezza sociale e protezione sociale;
- è necessario per la **tutela di interessi vitali** dell'interessato o di un'altra persona fisica, qualora l'interessato si trovi nell'incapacità fisica o giuridica di prestare consenso;
- è effettuato da **fondazioni, associazioni o altri organismi senza scopo di lucro** con finalità politiche, filosofiche, religiose o sindacali, nell'ambito delle loro attività legittime;
- riguarda dati resi **manifestamente pubblici** dall'interessato;
- è necessario per l'**accertamento, l'esercizio o la difesa di un diritto in sede giudiziaria**;
- è necessario per motivi di **interesse pubblico rilevante** sulla base del diritto dell'Unione o degli Stati membri;

---

<sup>264</sup> Regolamento (UE) 2016/679, art. 5.

- è necessario per finalità di **medicina preventiva o del lavoro**, valutazione della capacità lavorativa, diagnosi medica, assistenza o terapia sanitaria e sociale, sulla base del diritto o di un contratto con un professionista della sanità;
- è necessario per motivi di **interesse pubblico nel settore della sanità pubblica**, come la protezione da gravi minacce per la salute;
- è necessario per finalità di **archiviazione nel pubblico interesse, ricerca scientifica o storica o statistica**, nel rispetto delle garanzie previste dall'art. 89 del Regolamento.

Nel contesto lavorativo, è sicuramente rilevante la deroga che consente il trattamento quando esso è strettamente necessario per adempiere obblighi o esercitare diritti del datore di lavoro o del lavoratore in materia di diritto del lavoro, sicurezza e protezione sociale. In tale ipotesi, tuttavia, la liceità non è automatica: occorre che il trattamento sia proporzionato, giustificato da un'esigenza effettiva e accompagnato da garanzie adeguate, come misure di sicurezza rafforzate e limiti chiari alle finalità perseguite.<sup>265</sup>

Altre deroghe previste dall'art. 9, pur meno centrali rispetto al rapporto di lavoro, possono comunque avere un riflesso indiretto in questo ambito. Si pensi, ad esempio, ai trattamenti effettuati per finalità di medicina del lavoro e valutazione della capacità lavorativa, o ancora per motivi di sanità pubblica, come la prevenzione di malattie professionali o la protezione da minacce per la salute dei lavoratori. In questi casi, il trattamento di dati particolari si inserisce all'interno di un quadro normativo che bilancia esigenze collettive di sicurezza e tutela della salute con la protezione della dignità e della riservatezza individuale.

In definitiva, l'art. 9 opera un delicato bilanciamento: da un lato, la protezione rafforzata dei dati particolari come regola generale; dall'altro, la possibilità di deroghe giustificate da esigenze concrete e non surrogabili, che in ambito lavorativo trovano giustificazione soprattutto nelle materie della sicurezza, della protezione sociale e della salute del dipendente.

---

<sup>265</sup> Regolamento (UE) 2016/679, art. 9, par. 2, lett. b).

L'art. 88 GDPR<sup>266</sup> consente agli Stati membri di introdurre norme più specifiche al fine di garantire la protezione dei diritti e delle libertà dei lavoratori, soprattutto in relazione al trattamento dei dati personali, nel contesto del rapporto di lavoro. L'Italia ha recepito tale previsione nel **Codice Privacy** (D.lgs. 196/2003, come modificato dal D.lgs. 101/2018), introducendo limiti stringenti al trattamento di dati biometrici e sanitari: esso è consentito solo quando giustificato da esigenze effettive e non surrogabili con strumenti meno invasivi, e sempre nel rispetto dei principi di proporzionalità e necessità.<sup>267</sup>

Un'ulteriore tutela deriva dall'art 4 **dello Statuto dei lavoratori**, che disciplina l'uso di impianti e strumenti dai quali possa derivare un controllo a distanza sull'attività dei dipendenti. Tali strumenti possono essere installati ed utilizzati esclusivamente per esigenze organizzative e produttive, per la sicurezza del lavoro e per la tutela del patrimonio aziendale, e comunque previo accordo con le rappresentanze sindacali o, in mancanza, previa autorizzazione dell'Ispettorato del lavoro. Questi vincoli si applicano anche alle tecnologie basate su sistemi biometrici, che, per loro natura, possono comportare un monitoraggio costante e invasivo della persona.<sup>268</sup>

Come osserva Covelli, la trasparenza delle logiche algoritmiche è un requisito imprescindibile per salvaguardare la libertà e l'autodeterminazione del lavoratore. Comprendere in che modo i dati vengono raccolti, elaborati e trasformati in valutazioni o decisioni consente di contrastare fenomeni di opacità e bias sistemici che altrimenti rischierebbero di sfuggire a ogni controllo.<sup>269</sup> In questa prospettiva, il principio di minimizzazione impone di limitare il trattamento ai soli dati strettamente necessari, mentre il principio di proporzionalità richiede una valutazione comparativa sull'idoneità degli strumenti utilizzati, privilegiando sempre quelli meno invasivi quando siano disponibili alternative.

---

<sup>266</sup> Regolamento (UE) 2016/679, art. 88.

<sup>267</sup> D.lgs. 30 giugno 2003, n. 196, *Codice in materia di protezione dei dati personali*, art. 113, come modificato dal D.lgs. 10 agosto 2018, n. 101.

<sup>268</sup> Legge 20 maggio 1970, n. 300, *Statuto dei Lavoratori*, art. 4.

<sup>269</sup> DONINI, A. M., COVELLI, R., *Lavoro e intelligenza artificiale: dalla trasparenza alla conoscibilità*, in *Labour & Law Issues*, vol. 9, n. 1, 2023, p. 91.

L'analisi di questi vincoli normativi e dei relativi meccanismi di tutela che ne discendono, rende necessario soffermarsi sul quadro legislativo nazionale ed europeo applicabile, che costituirà l'oggetto del paragrafo successivo.

## 5. Controllo algoritmico e autonomia decisionale del lavoratore

L'avvento delle piattaforme digitali e dei sistemi di intelligenza artificiale ha rivoluzionato radicalmente le modalità di esercizio del potere direttivo e di controllo nelle organizzazioni lavorative. Se in passato la subordinazione si manifestava prevalentemente attraverso ordini diretti e immediati del datore di lavoro, oggi essa assume i tratti di una "subordinazione algoritmica", in cui è il software a stabilire tempi, modalità e criteri della prestazione.

Le piattaforme digitali, da Uber a Deliveroo, fino alle più recenti applicazioni di smart management in contesti aziendali tradizionali, rappresentano esempi paradigmatici di questa trasformazione. Il lavoratore non riceve più direttive da un superiore gerarchico visibile, ma da un "capo invisibile", incarnato dall'algoritmo che assegna gli incarichi, stabilisce i tempi di consegna, calcola le retribuzioni e, nei casi più estremi, arriva persino a disattivare unilateralmente l'account del prestatore. Questa forma di eterodirezione digitale genera inevitabilmente un rapporto caratterizzato da opacità, automatismo e apparente neutralità tecnica, ridefinendo la stessa configurazione del potere datoriale.

La giurisprudenza internazionale ha iniziato a riconoscere il fenomeno. Emblematico è il caso deciso l'11 aprile 2022 del Tribunal Superior do Trabalho del Brasile<sup>270</sup>, che per la prima volta ha qualificato il rapporto tra un autista Uber e la piattaforma come lavoro subordinato. Nella motivazione, la Corte ha affiancato alla subordinazione classica, la nozione di subordinazione algoritmica, intesa come il controllo esercitato mediante sistemi sofisticati di raccolta dati, intelligenza artificiale e strumenti telematici in grado di dirigere e disciplinare, in modo pervasivo, l'attività del lavoratore<sup>271</sup>. La pronuncia ha

---

<sup>270</sup> TRIBUNAL SUPERIOR DO TRABALHO (Brasile), 3ª Turma, *Recurso de Revista* n. 100353-02.2017.5.01.0066, 11 aprile 2022.

<sup>271</sup> LAUANDE RODRIGUES, P., *Trabalho em plataformas digitais, subordinação algorítmica e presunção do vínculo empregatício*, in *Lavoro e Previdenza Oggi*, n. 9–10, 2022, p. 659 ss.

sottolineato come la piattaforma non si limitasse a mettere in contatto conducenti e passeggeri, ma svolgesse una vera e propria attività imprenditoriale fondata sulla gestione algoritmica della forza lavoro, richiamando esplicitamente il principio di “primazia della realtà” tipico del diritto del lavoro brasiliano.<sup>272</sup>

La dottrina ha colto con chiarezza questa metamorfosi, osservando come il management algoritmico dia vita ad una nuova dimensione della subordinazione: i lavoratori sono soggetti a valutazioni continue e a regole organizzative che limitano fortemente la loro autonomia decisionale.<sup>273</sup> Non si tratta solo di misurare l’efficienza individuale, ma di incidere direttamente sulle opportunità di accesso al mercato, prevedendo premi, penalizzazioni ed esclusioni dalla piattaforma in caso di performance ritenute insufficienti.<sup>274</sup>

Parallelamente, anche l’ordinamento europeo ha iniziato a reagire. La recente **Direttiva (UE) sul lavoro mediante piattaforme digitali**, approvata dal Parlamento Europeo il 24 aprile 2024, introduce all’art. 5 una presunzione relativa di subordinazione: ogniqualvolta emergano elementi che indicano un potere di controllo o di direzione da parte della piattaforma, il rapporto deve presumersi di natura subordinata, salvo prova contraria fornita dal committente<sup>275</sup>. Tale formulazione, frutto di un difficile compromesso tra le istituzioni europee, segna un passo avanti rispetto alla proposta originaria della Commissione del 2021, che subordinava l’operatività della presunzione alla dimostrazione della presenza di almeno due dei cinque indici tipizzati (determinazione unilaterale della retribuzione, supervisione elettronica, limitazione della libertà di scelta degli incarichi, ecc.).<sup>276</sup>

---

<sup>272</sup> Ivi, p. 662.

<sup>273</sup> ADAMS-PRASSL, J., *What if your boss was an algorithm? Economic incentives, legal challenges, and the rise of artificial intelligence at work*, in *Comparative Labor Law & Policy Journal*, vol. 41, 2019, p. 123 ss.

<sup>274</sup> MOORE, P. V., UPCHURCH, M., WHITTAKER, X. (a cura di), *Humans and machines at work*, Palgrave Macmillan, Cham, 2018, p. 17 ss.

<sup>275</sup> BIASI, M., *La presunzione relativa di subordinazione nella “Direttiva Piattaforme”*: *brevi spunti*, in *Lavoro Diritti Europa*, n. 4, 2024, p. 3.

<sup>276</sup> COMMISSIONE EUROPEA, *Proposta di Direttiva sul miglioramento delle condizioni di lavoro dei lavoratori delle piattaforme*, COM(2021) 762 final, 9 dicembre 2021, art. 4, par. 2.

Pur rappresentando un progresso, la scelta di una presunzione relativa lascia ampi margini interpretativi agli Stati membri, con il rischio di produrre soluzioni disomogenee a livello europeo. Nel contesto italiano, tuttavia la portata innovativa della direttiva appare più contenuta. Considerando che già l'art. 2094 c.c. e l'art. 2 del d.lgs. 81/2015 consentono di qualificare rapporti etero-diretti come subordinati o "etero-organizzati", la portata innovativa della Direttiva appare ridotta. Come osserva Biasi, la disposizione europea rischia di essere "sostanzialmente inutile" nel nostro ordinamento, poiché gli stessi elementi che la Direttiva qualifica come indici presuntivi costituiscono già tratti discretivi della subordinazione nel diritto interno<sup>277</sup>. Tuttavia, la sua valenza simbolica e politica resta significativa: la Direttiva segna il riconoscimento, a livello sovranazionale, che il potere algoritmico è una forma di direzione e controllo giuridicamente rilevante, con implicazioni decisive per la tutela dei diritti dei lavoratori.<sup>278</sup>

In questo scenario, la perdita di autonomia decisionale non si traduce soltanto in un fenomeno individuale, ma investe la struttura stessa del rapporto di lavoro. La figura del "capo invisibile" mette in discussione le categorie tradizionali del diritto del lavoro, imponendo una riflessione sulle nuove tutele individuali e collettive necessarie per bilanciare l'opacità e la pervasività del controllo algoritmico.

Se il management algoritmico rappresenta la forma più pervasiva di eterodirezione digitale, lo smart working costituisce invece il volto più flessibile e apparentemente emancipante delle trasformazioni indotte dalla digitalizzazione, ponendo al centro nuove modalità di organizzazione della prestazione lavorativa.

## 6. Lo "smart working"

L'adozione dello "smart working", o lavoro agile, si inserisce nel più ampio quadro delle trasformazioni indotte dalla digitalizzazione e dalla crescente flessibilità organizzativa delle imprese. Pur conservando la natura subordinata del rapporto di lavoro, esso segna una netta discontinuità rispetto al telelavoro tradizionale. Come evidenziato dalla dottrina, il lavoro agile si distingue per l'assenza di una postazione fissa e per

---

<sup>277</sup> BIASI, M., op. cit., p. 6.

<sup>278</sup> Ivi, p. 8 ss.

l'organizzazione delle attività per fasi, cicli e obiettivi, superando i vincoli spaziali e temporali tipici dell'attività lavorativa d'ufficio.<sup>279</sup> Tale modello privilegia il raggiungimento dei risultati, riconoscendo al lavoratore un'ampia autonomia nella scelta del luogo di lavoro e nella gestione dei tempi, pur nel rispetto degli obiettivi e delle modalità concordate con il datore di lavoro.<sup>280</sup>

Esso si svolge in parte all'interno dei locali aziendali e in parte all'esterno, senza postazione fissa e con l'ausilio di strumenti tecnologici idonei.<sup>281</sup> L'accordo individuale tra datore e lavoratore costituisce il fulcro regolativo, attraverso il quale vengono disciplinati tempi di lavoro, forme di esercizio del potere direttivo e criteri di utilizzo degli strumenti informatici.<sup>282</sup>

In Italia la diffusione dello smart working ha seguito inizialmente una fase graduale di sperimentazione, limitata a contesti aziendali innovativi, fino a conoscere una repentina accelerazione durante la pandemia da Covid-19. L'emergenza sanitaria ne ha infatti imposto un ricorso massivo, spesso in assenza di una pianificazione organizzativa adeguata e di strumenti idonei alla valutazione dei rischi, sollevando questioni estremamente delicate sia sul piano giuridico sia su quello operativo.<sup>283</sup>

Accanto agli enormi benefici, quali sicuramente una maggiore conciliazione vita-lavoro, la riduzione degli spostamenti e il potenziale incremento della produttività, il lavoro agile pone una serie di **criticità rilevanti e complesse**: la diluizione dei confini tra sfera privata e professionale, il rischio di isolamento sociale, la gestione del carico di lavoro e della reperibilità continua, nonché la necessità di garantire condizioni di salute e sicurezza equivalenti a quelle del lavoro in sede.<sup>284</sup> A tali sfide si aggiunge l'impatto della tecnologia sui meccanismi di controllo e valutazione delle performance, con implicazioni

---

<sup>279</sup>GENTILINI, V., FILOSA, R., *La tutela della salute e sicurezza del lavoro nello smart working*, Working Paper n. 20, p. 4.

<sup>280</sup>AA. VV., *Il futuro del lavoro si chiama "smart working"? Riflessioni e prospettive*, in *Rivista trimestrale di Organizzazione Aziendale*, p. 2.

<sup>281</sup> GENTILINI, V., FILOSA, op. cit., p. 5.

<sup>282</sup> Ibidem.

<sup>283</sup> *Rivista trimestrale di Organizzazione Aziendale*, *Il futuro del lavoro si chiama "smart working"?*, cit., p. 3.

<sup>284</sup> GENTILINI, V., FILOSA, op. cit., p. 6.

dirette sulla privacy del lavoratore e sull'equilibrio dei rapporti di potere all'interno dell'organizzazione.<sup>285</sup>

## 6.1 Inquadramento normativo e sfide di tutela

La **L.81/2017** rappresenta il riferimento primario in materia di smart working, stabilendo, all'art. 22, che al lavoratore agile si applichino le stesse tutele previste per i lavoratori che operano all'interno dei locali aziendali in materia di salute e sicurezza.<sup>286</sup> In particolare, il datore di lavoro è tenuto a fornire al lavoratore e al rappresentante dei lavoratori per la sicurezza (RLS) un'informativa scritta sui rischi generali e specifici connessi alla modalità di esecuzione della prestazione.<sup>287</sup> Tale obbligo implica la valutazione delle condizioni ambientali in cui si svolge l'attività, nonché delle caratteristiche delle attrezzature e degli strumenti utilizzati.<sup>288</sup>

Il lavoratore agile, a sua volta, deve cooperare all'attuazione delle misure di prevenzione predisposte dal datore di lavoro, rispettando le istruzioni ricevute e le norme vigenti in materia di sicurezza.<sup>289</sup> Restano invariati gli obblighi generali di formazione e di sorveglianza sanitaria, che devono però essere adattati alle specificità del lavoro da remoto.<sup>290</sup>

Tra i profili più discussi vi è quello del diritto alla disconnessione, introdotto per prevenire i rischi connessi all'iperconnessione e all'invasione della sfera privata da parte delle esigenze lavorative.<sup>291</sup> Questo diritto, richiamato anche in diverse intese aziendali e contratti collettivi, si configura come uno strumento essenziale per la prevenzione dei rischi psicosociali e dello stress lavoro-correlato.

---

<sup>285</sup>Rivista trimestrale di Organizzazione Aziendale, *Il futuro del lavoro si chiama "smart working"?*, cit., p. 4.

<sup>286</sup> GENTILINI, V., FILOSA, op. cit., p. 8.

<sup>287</sup> Ibidem.

<sup>288</sup> GENTILINI, V., FILOSA, op. cit., p. 9.

<sup>289</sup> Ibidem.

<sup>290</sup> Ivi., p. 10.

<sup>291</sup> Rivista trimestrale di Organizzazione Aziendale, *Il futuro del lavoro si chiama "smart working"?*, cit., p. 5

L'emersione del diritto alla disconnessione rappresenta una delle sfide più significative poste dal lavoro digitale e dallo smart working. La possibilità tecnica di essere costantemente reperibili tramite strumenti informatici e reti aziendali ha infatti favorito una progressiva erosione dei confini tra tempo di lavoro e tempo di vita, con effetti rilevanti sul benessere psicofisico dei lavoratori e sul loro equilibrio familiare e sociale. La dottrina ha sottolineato come il rischio maggiore sia quello di una "iper-connessione" permanente, che trasforma la flessibilità promessa dal lavoro agile in una nuova forma di subordinazione occulta, basata sulla disponibilità continua del prestatore di lavoro.<sup>292</sup>

Un primo riferimento normativo espresso al diritto alla disconnessione è stato introdotto proprio con la L. n. 81/2017,<sup>293</sup> dedicata al lavoro agile. L'art. 19, comma 1, prevede che l'accordo individuale disciplini "i tempi di riposo del lavoratore" e "le misure tecniche e organizzative necessarie per assicurare la disconnessione" dagli strumenti tecnologici di lavoro. Tuttavia, tale formulazione lascia ampi margini di discrezionalità, affidando la concreta tutela a una negoziazione caso per caso, senza individuare standard uniformi e vincolanti.

Sul piano europeo, il dibattito si è intensificato negli ultimi anni. La Direttiva (UE) 2019/1152<sup>294</sup>, pur non menzionando espressamente la disconnessione, ha rafforzato il principio della trasparenza e della prevedibilità delle condizioni di lavoro, richiedendo che gli Stati membri garantiscano ai lavoratori informazioni chiare sui tempi e sulle modalità della prestazione. Alcuni ordinamenti hanno però compiuto passi più decisi. In Francia, con la Loi n. 1088/2016<sup>295</sup>, le imprese con più di cinquanta dipendenti sono

---

<sup>292</sup> RICCI, A., *Il diritto alla disconnessione tra benessere del lavoratore e nuove forme di subordinazione digitale*, in *Rivista italiana di diritto del lavoro*, 2022, p. 15.

<sup>293</sup> Legge 22 maggio 2017, n. 81, *Misure per la tutela del lavoro autonomo non imprenditoriale e misure volte a favorire l'articolazione flessibile nei tempi e nei luoghi del lavoro subordinato* (c.d. Jobs Act del lavoro autonomo e sul lavoro agile), in *Gazzetta Ufficiale* n. 135 del 13 giugno 2017, art. 19.

<sup>294</sup> Direttiva (UE) 2019/1152 del Parlamento europeo e del Consiglio del 20 giugno 2019, relativa a condizioni di lavoro trasparenti e prevedibili nell'Unione europea, in *Gazzetta ufficiale dell'Unione europea* (GUUE) L 186 dell'11 luglio 2019.

<sup>295</sup> Loi n° 2016-1088 du 8 août 2016, relative au travail, à la modernisation du dialogue social et à la sécurisation des parcours professionnels, in *Journal officiel de la République française* (JORF), n. 0184 del 9 août 2016.

obbligate a disciplinare la disconnessione nei contratti collettivi.<sup>296</sup> La Spagna ha introdotto un diritto esplicito alla disconnessione digitale con la Ley Orgánica 3/2018<sup>297</sup>, che tutela la riservatezza e garantisce il riposo dei lavoratori anche nel contesto delle comunicazioni aziendali digitali. In Belgio, una circolare ministeriale del 2018<sup>298</sup> ha incoraggiato il dialogo sociale nelle imprese sul tema, mirando a definire soluzioni condivise per prevenire l'abuso della reperibilità costante.<sup>299</sup>

Queste esperienze mostrano come il diritto alla disconnessione possa assumere configurazioni diverse: da obblighi negoziali collettivi (Francia), a diritti soggettivi direttamente azionabili (Spagna), fino a politiche di soft law che promuovono la contrattazione (Belgio). In tutti i casi, tuttavia, emerge la consapevolezza che la disconnessione non può essere affidata alla sola autonomia individuale, ma necessita di cornici normative o collettive capaci di riequilibrare il rapporto di forze tra datore di lavoro e lavoratore.<sup>300</sup>

La giurisprudenza e la dottrina italiane hanno inoltre posto in rilievo il collegamento del diritto alla disconnessione con principi costituzionali fondamentali. Esso si ricollega, da un lato, all'art. 36 Cost.<sup>301</sup>, che garantisce il diritto al riposo, e dall'altro all'art. 32 Cost.<sup>302</sup>, che tutela la salute come diritto inviolabile dell'individuo. Alcuni autori hanno interpretato il diritto alla disconnessione come presidio della dignità della persona, sottolineando che il lavoratore deve poter esercitare un controllo effettivo sui propri tempi di vita e di lavoro, senza essere intrappolato in un flusso comunicativo costante e potenzialmente lesivo della sua integrità psico-fisica.<sup>303</sup>

---

<sup>296</sup> THALION, M., *The right to disconnect: work-life balance and digital labour law in Europe*, in *International Journal of Comparative Labour Law*, 2021.

<sup>297</sup> *Ley Orgánica 3/2018, de 5 de diciembre*, de Protección de Datos Personales y garantía de los derechos digitales, in *BOE n. 294 del 6 diciembre 2018*.

<sup>298</sup> *Loi du 26 mars 2018, relative au renforcement de la croissance économique et de la cohésion sociale*, artt. 15-17, in *Moniteur belge*, 30 mars 2018.

<sup>299</sup> THALION, M., op. cit.

<sup>300</sup> *Ibidem*.

<sup>301</sup> Art. 36 Cost., in G.U. 27 dicembre 1947, n. 298, ed. straord.

<sup>302</sup> Art. 32 Cost., in G.U. 27 dicembre 1947, n. 298, ed. straord.

<sup>303</sup> RICCI, A., op. cit., p. 20.

In definitiva, il diritto alla disconnessione si configura come un istituto in via di consolidamento, al crocevia tra diritto del lavoro, salute e nuove tecnologie. Esso costituisce una risposta necessaria alle trasformazioni del lavoro agile, ma al tempo stesso apre la strada a un ripensamento complessivo delle tutele, nella prospettiva di un diritto del lavoro sempre più orientato alla salvaguardia della persona del lavoratore anche nello spazio digitale. Non sorprende, quindi, che la disconnessione venga oggi discussa non solo come misura organizzativa, ma come possibile diritto fondamentale di nuova generazione, destinato a incidere sul futuro assetto del diritto del lavoro in Europa.<sup>304</sup>

Un'ulteriore questione riguarda il perimetro dei controlli a distanza: l'uso di strumenti tecnologici per monitorare l'attività del lavoratore deve rispettare i limiti fissati dall'art. 4 dello Statuto dei lavoratori, il quale consente controlli solo per esigenze organizzative, produttive o di sicurezza, previa adeguata informativa e con garanzie di trasparenza.<sup>305</sup>

Infine, un elemento cruciale emerso dalla prassi è la formazione a distanza (FAD) in materia di salute e sicurezza. Essa offre vantaggi quali l'accessibilità e la riduzione dei costi, ma richiede l'adozione di metodologie interattive e personalizzate per garantire l'efficacia, in particolare per la prevenzione dei rischi specifici legati allo smart working.<sup>306</sup> In questo contesto, la contrattazione collettiva può svolgere un ruolo strategico nel fissare standard minimi di tutela e nel promuovere buone pratiche organizzative.<sup>307</sup>

## **7. Normativa nazionale ed europea applicabile**

L'introduzione dell'intelligenza artificiale nei processi lavorativi hanno reso necessario un ripensamento della cornice normativa tradizionale, sia a livello europeo che nazionale. La disciplina giuridica, infatti, non si limita a regolare il trattamento dei dati personali, ma si estende alla definizione di limiti, diritti e garanzie specifiche che tutelino la dignità. Se il diritto del lavoro classico si fondava su una dimensione prevalentemente "analogica" del rapporto tra datore e prestatore, la progressiva datificazione dell'attività lavorativa ha

---

<sup>304</sup> THALION, M., op. cit.

<sup>305</sup> GENTILINI, V., FILOSA, R., op. cit., p. 11.

<sup>306</sup> Ivi, p. 12.

<sup>307</sup> Ivi, p. 13

spostato l'asse verso modelli organizzativi incentrati sulla sorveglianza digitale e sulla valutazione automatizzata. In questo contesto, la coesistenza di norme di portata generale, come il **Regolamento (UE) 2016/679** (*General Data Protection Regulation – GDPR*), e di discipline settoriali, quali lo **Statuto dei lavoratori** e il **Codice Privacy**, deve essere integrata da nuove fonti, come il Regolamento europeo sull'IA (*AI Act*), che introducono obblighi di trasparenza, controllo e responsabilità specifici per i sistemi algoritmici.

## 5.1 Il quadro normativo europeo

L'Unione europea sta delineando un sistema regolatorio complesso per disciplinare l'impiego dell'intelligenza artificiale nei processi lavorativi, fondato su un intreccio di norme generali e settoriali. Due pilastri principali sorreggono questo quadro: **il Regolamento sull'Intelligenza Artificiale (AI Act)**, ancora in fase di approvazione ma già definito nei suoi contenuti essenziali, e la **Direttiva (UE) 2024/2831** sul lavoro tramite piattaforme digitali. Entrambe si coordinano con la disciplina generale sulla protezione dei dati personali prevista dal **Regolamento (UE) 2016/679 (GDPR)**, integrando e rafforzando le garanzie a tutela dei lavoratori.<sup>308</sup>

La proposta di AI Act adotta un approccio basato sulla classificazione dei sistemi di intelligenza artificiale in quattro categorie di rischio: inaccettabile, alto, limitato e minimo<sup>309</sup>. I sistemi utilizzati per il reclutamento, la selezione, la valutazione o l'assegnazione delle mansioni rientrano nella categoria ad "alto rischio" in base all'art. 6 e all'Allegato III, punto 4, poiché incidono in modo significativo sui diritti fondamentali e sulle opportunità professionali.<sup>310</sup> Per questi sistemi, il regolamento prevede requisiti stringenti che comprendono la predisposizione di procedure di gestione del rischio mediante valutazioni preventive e misure di mitigazione<sup>311</sup> la garanzia di qualità dei dati di addestramento, validazione e test, i quali devono essere pertinenti, rappresentativi e

---

<sup>308</sup> COVELLI, A., *Privacy e lavoro: la gestione algoritmica dei rapporti lavorativi*, in COVELLI, R., *Lavoro e intelligenza artificiale: dai principi di trasparenza algoritmica al diritto alla conoscibilità*, Università degli Studi di Milano-Bicocca, in *LLI – Labour & Law Issues*, vol. 9, n. 1, 2023, p. 12.

<sup>309</sup> Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (AI Act), art. 5.

<sup>310</sup> Ivi, art. 6 e Allegato III, punto 4.

<sup>311</sup> Ivi, art. 9.

privi di *bias* ingiustificati<sup>312</sup>; la redazione di documentazione tecnica e registri che consentano la tracciabilità delle decisioni algoritmiche<sup>313</sup>; l'obbligo di fornire agli utenti informazioni chiare sulle capacità e sui limiti del sistema<sup>314</sup>; e la previsione di un'adeguata sorveglianza umana per prevenire effetti pregiudizievoli e garantire la possibilità di intervento<sup>315</sup>. Il considerando 36 del regolamento sottolinea inoltre che l'uso dell'IA in ambito lavorativo deve rispettare i principi di proporzionalità e necessità, evitando trattamenti sproporzionati rispetto alle finalità perseguite.<sup>316</sup>

Parallelamente, la Direttiva (UE) 2024/2831 interviene in maniera mirata sui meccanismi di gestione algoritmica del lavoro mediato da piattaforme, fornendo un modello che può fungere da parametro interpretativo anche per altre forme di organizzazione aziendale basate su strumenti di monitoraggio e valutazione automatizzata.<sup>317</sup> La direttiva stabilisce obblighi informativi dettagliati riguardo al funzionamento dei sistemi algoritmici, alle finalità del trattamento e al loro impatto sulle condizioni di lavoro (art. 6)<sup>318</sup>; vieta il trattamento di dati sensibili, come quelli relativi alla salute, alle opinioni politiche o all'appartenenza sindacale (art. 7)<sup>319</sup>; impone che le decisioni automatizzate aventi effetti giuridici o significativi siano sottoposte a supervisione e revisione umana (art. 8)<sup>320</sup>; riconosce ai lavoratori il diritto di chiedere il riesame delle decisioni algoritmiche (art. 9)<sup>321</sup>; e garantisce ai rappresentanti dei lavoratori l'accesso alle informazioni necessarie

---

<sup>312</sup> Ivi, art. 10

<sup>313</sup> Ivi, art. 11.

<sup>314</sup> Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (AI Act), art. 13

<sup>315</sup> Ivi, art. 14

<sup>316</sup> Ivi, considerando 36.

<sup>317</sup> Direttiva (UE) 2024/2831 del Parlamento europeo e del Consiglio del 13 dicembre 2024, relativa al miglioramento delle condizioni di lavoro nel lavoro mediante piattaforme digitali, in *Gazzetta ufficiale dell'Unione europea* (GUUE), L 202 dell'8 gennaio 2025, p. 17, art. 6.

<sup>318</sup> Ivi, art. 6.

<sup>319</sup> Ivi, art. 7.

<sup>320</sup> Ivi, art. 8.

<sup>321</sup> Ivi, art. 9.

per monitorare il rispetto delle norme (art. 10).<sup>322</sup> Tale approccio mira a rafforzare la trasparenza e ad assicurare una tutela effettiva contro gli automatismi opachi.

Questo impianto normativo si integra con le disposizioni generali del **Regolamento (UE) 2016/679 (GDPR)**, di cui nel paragrafo precedente sono stati richiamati i principi fondamentali. Basti qui ricordare come L'art. 5 stabilisce i principi di liceità, correttezza, trasparenza, minimizzazione, esattezza e sicurezza dei dati<sup>323</sup>; l'art. 9 vieta in via generale il trattamento di categorie particolari di dati, tra cui quelli biometrici e sanitari, salvo specifiche eccezioni<sup>324</sup>; mentre l'art. 88 consente agli Stati membri di introdurre norme più specifiche per garantire la protezione dei lavoratori.<sup>325</sup> Come osserva Covelli, la trasparenza delle logiche algoritmiche e l'accesso alle informazioni costituiscono condizioni imprescindibili per tutelare l'autodeterminazione del lavoratore e contrastare fenomeni di discriminazione o distorsione sistemica.<sup>326</sup>

## 5.2 Il quadro normativo nazionale

L'ordinamento italiano affronta il tema del controllo datoriale e del trattamento dei dati personali dei lavoratori con un impianto normativo molto complesso, volto a bilanciare due esigenze solo in apparenza contrapposte: da un lato, la tutela della dignità e della riservatezza dei dipendenti; dall'altro, le esigenze organizzative e produttive dell'impresa. Tale disciplina si fonda principalmente sullo **Statuto dei lavoratori** (L. 20 maggio 1970, n. 300), sul **Codice in materia di protezione dei dati personali** (D.lgs. 30 giugno 2003, n. 196, come modificato dal D.lgs. 10 agosto 2018, n. 101 per l'adeguamento al GDPR) e sui provvedimenti del **Garante per la protezione dei dati personali**.

---

<sup>322</sup> Ivi, art. 10

<sup>323</sup> Regolamento (UE) 2016/679 (GDPR), art. 5

<sup>324</sup> Ivi, art. 9.

<sup>325</sup> Ivi, art. 88.

<sup>326</sup> COVELLI, R., *Lavoro e intelligenza artificiale: dalla trasparenza alla conoscibilità*, in *Labour & Law Issues*, vol. 9, n. 1, 2023, p. 91.

Uno dei cardini di questa regolazione rimane **l'art. 4 dello Statuto dei lavoratori**<sup>327</sup>, già richiamato nella parte dedicata alla privacy come norma di garanzia rispetto ai rischi di controllo invasivo. Qui è opportuno soffermarsi nuovamente su di esso perché, alla luce dell'introduzione di sistemi di intelligenza artificiale, la disposizione acquista una valenza ulteriore: non solo limite al trattamento dei dati e alla sorveglianza a distanza, ma vero e proprio strumento di **bilanciamento tra potere organizzativo del datore e dignità del prestatore**. La riforma del 2015 (D.lgs. 151/2015) ha reso esplicite le finalità che legittimano l'uso di impianti e strumenti di controllo, esigenze produttive, sicurezza sul lavoro e tutela del patrimonio aziendale, subordinandolo a un accordo sindacale o, in mancanza, all'autorizzazione dell'Ispettorato nazionale del lavoro.

Se nel discorso sulla privacy l'art. 4 fungeva da barriera contro ingerenze indebite, nel contesto della **gestione algoritmica** si impone come parametro di **liceità e proporzionalità** delle nuove tecnologie, specie quelle idonee a realizzare sorveglianza biometrica o monitoraggio continuo. La "recitazione" di questa norma non è quindi mera ripetizione, ma il riconoscimento del suo ruolo di clausola generale, capace di adattarsi a fenomeni che nel 1970 erano impensabili, ma che oggi incidono profondamente sull'equilibrio del rapporto di lavoro.

A questa disciplina si affianca il **Codice Privacy**, che agli artt.113<sup>328</sup> e 114<sup>329</sup> richiama espressamente l'obbligo di conformarsi nel trattamento dei dati in ambito lavorativo ai limiti fissati dallo Statuto. In attuazione dell'art. 88 GDPR, inoltre, il legislatore italiano, ha introdotto disposizioni specifiche volte a rafforzare la protezione dei dati dei lavoratori, con particolare attenzione ai trattamenti di dati sensibili, tra cui quelli biometrici e sanitari. Tali trattamenti sono consentiti solo se strettamente necessari per adempiere

---

<sup>327</sup> Art. 4, Legge 20 maggio 1970, n. 300, *Statuto dei lavoratori*, in *Gazzetta Ufficiale* 27 maggio 1970, n. 131, come modificato dal D.lgs. 14 settembre 2015, n. 151, in *Gazzetta Ufficiale* 23 settembre 2015, n. 221.

<sup>328</sup> Art. 113, D.lgs. 30 giugno 2003, n. 196, *Codice in materia di protezione dei dati personali*, come modificato dal D.lgs. 10 agosto 2018, n. 101.

<sup>329</sup> Art. 114, D.lgs. 196/2003.

obblighi di legge o di contratto collettivo, o per esercitare diritti in materia di diritto del lavoro, sicurezza e protezione sociale, e sempre nel rispetto di garanzie adeguate.<sup>330</sup>

Il **Garante per la protezione dei dati personali** ha più volte sottolineato che l'adozione di sistemi tecnologici in ambito lavorativo deve essere preceduta da una valutazione di impatto sulla protezione dei dati (DPIA), ai sensi dell'art. 35 GDPR, soprattutto quando tali sistemi comportano un monitoraggio sistematico e su larga scala dei lavoratori.<sup>331</sup> Nel caso di utilizzo di strumenti basati su IA, tale valutazione deve considerare anche i rischi di *bias* algoritmico e di decisioni automatizzate, in linea con l'art. 22 GDPR, assicurando sempre un'adeguata supervisione umana.

Sul piano pratico, la disciplina nazionale tende a bilanciare il potere organizzativo e direttivo del datore con il diritto del lavoratore a non subire ingerenze indebite nella propria sfera privata. Come osserva Marazza, ciò implica non soltanto il rispetto formale delle condizioni di liceità, ma anche un'applicazione sostanziale dei principi di minimizzazione e proporzionalità, valutando se l'obiettivo aziendale possa essere raggiunto con strumenti meno invasivi.<sup>332</sup> In questa prospettiva, il diritto del lavoro italiano dimostra una sensibilità crescente verso le sfide introdotte dall'automazione e dalla gestione algoritmica, pur lasciando emergere la necessità di interventi normativi aggiornati che tengano conto delle peculiarità dei sistemi di intelligenza artificiale e della loro capacità predittiva e pervasiva.

Il confronto con il livello europeo mette in evidenza un nodo centrale: la regolazione dell'intelligenza artificiale nei processi lavorativi non può ridursi a un mero adempimento formale, ma deve assumere la forma di una **tutela sostanziale**, fondata su principi comuni e condivisi. La cornice europea – attraverso l'**AI Act**, la **Direttiva sul lavoro tramite piattaforme digitali e il GDPR** – stabilisce una base uniforme che valorizza la valutazione preventiva del rischio, la trasparenza delle logiche algoritmiche e la necessità

---

<sup>330</sup> Art. 9, par. 2, lett. b) e h), Regolamento (UE) 2016/679 (GDPR); art. 2-septies, D.lgs. 196/2003.

<sup>331</sup> GARANTE PER LA PROTEZIONE DEI DATI PERSONALI, *Provvedimento recante "Linee guida sui trattamenti di dati personali per finalità di gestione del rapporto di lavoro in ambito pubblico e privato"*, Roma, 5 giugno 2019, doc. web n. 9124510.

<sup>332</sup> MARAZZA, M., *Dei poteri (del datore di lavoro), dei controlli (a distanza) e del trattamento dei dati (del lavoratore)*, in *Biblioteca 20 Maggio*, n. 1, 2016, p. 3

di una supervisione umana effettiva. Parallelamente, la normativa italiana – dal **Codice Privacy** allo **Statuto dei lavoratori** – introduce limiti più rigorosi in materia di trattamento dei dati sensibili e di controlli a distanza, rafforzando in particolare la protezione della dignità e della libertà del lavoratore.

Da questa interazione multilivello scaturisce un **sistema di garanzie composito**, che consente di affrontare in maniera coordinata le sfide poste dall'uso dell'intelligenza artificiale nel contesto lavorativo. L'efficacia di tale sistema, tuttavia, dipende dalla capacità dei datori di lavoro di interpretare e applicare in modo integrato le diverse fonti normative, evitando approcci frammentari e garantendo che l'innovazione tecnologica sia sempre accompagnata da una valutazione preventiva del suo impatto sui diritti fondamentali.

## CAPITOLO 3

### Un'Intelligenza Artificiale per l'Inclusione: Il Progetto “Engage” a Tilburg

#### Introduzione

Nell'attuale scenario europeo, la transizione digitale pone sfide nuove e complesse, non solo sotto il profilo tecnologico, ma anche in termini di equità sociale ed inclusione. L'automazione, l'uso diffuso dell'intelligenza artificiale e l'evoluzione delle competenze richieste dal mercato del lavoro rischiano infatti di amplificare le disuguaglianze esistenti, colpendo le fasce più vulnerabili della popolazione. In assenza di una riflessione etica condivisa e di un orientamento strategico verso la sostenibilità e l'umanità dei processi innovativi, il progresso tecnologico rischia di tradursi in un ulteriore motivo di esclusione sociale.

È in questo nuovo e difficile scenario che si inserisce il progetto elaborato nell'ambito del programma *ENGAGE* presso l'università di Tilburg, oggetto di approfondimento del presente capitolo. Si tratta di un caso studio concreto che incarna l'applicazione pratica dei principi fin qui esaminati. Ho avuto l'opportunità di prenderne parte in rappresentanza dell'università LUISS Guido Carli, contribuendo attivamente alla definizione dell'idea progettuale ed alla sua realizzazione. L'esperienza si è conclusa con la vittoria del nostro gruppo alla challenge finale, attestando la solidità di una proposta che ha saputo coniugare innovazione tecnologica, inclusione sociale e centralità della persona.

La visione che ha ispirato il lavoro progettuale riflette pienamente i principi promossi dalla professoressa Laura Pallazzani, con la quale ho avuto l'onore di collaborare. Al centro dell'iniziativa vi è la convinzione che l'intelligenza artificiale debba porsi al servizio dell'essere umano, rispettandone la dignità, valorizzandone le capacità e promuovendo l'inclusione attiva di soggetti spesso esclusi dai processi produttivi e decisionali.

Il progetto ha avuto come obiettivo la creazione di una piattaforma digitale, ispirata ad un paradigma antropocentrico, finalizzata a facilitare l'inserimento lavorativo di individui

con background migratorio e di altri gruppi vulnerabili, attraverso l'integrazione di strumenti di design thinking, principi di etica dell'IA e soluzioni tecnologiche avanzate. Il presente capitolo si propone di analizzare in modo approfondito le diverse fasi del progetto, le metodologie adottate ed infine i risultati conseguiti, evidenziandone il rilievo sotto il profilo etico, giuridico e sociale.

Attraverso un approccio sistemico, partecipativo e interdisciplinare, il gruppo di lavoro ha esplorato le potenzialità dell'intelligenza artificiale non come sostituto della relazione umana, ma come catalizzatore di processi inclusivi, trasparenti ed orientati al bene comune. L'IA viene reinterpretata come strumento per rafforzare le capacità individuali, promuovere l'autonomia e contribuire allo sviluppo di modelli di governance equi e replicabili, capaci di adattarsi a diversi contesti territoriali e culturali.

## **1. Contesto e obiettivo del progetto**

Il progetto oggetto di analisi è stato sviluppato nel quadro dell'iniziativa *Engage*, un programma europeo, volto a connettere giovani professionisti e studenti universitari provenienti da differenti paesi e ambiti disciplinari, con l'obiettivo ultimo di promuovere soluzioni innovative alle grandi sfide sociali del nostro tempo. L'iniziativa si basa su un approccio collaborativo e partecipativo, in cui il dialogo interculturale e l'interazione tra conoscenze diverse costituiscono il fulcro metodologico per l'elaborazione di proposte definitive a queste rilevanti problematiche attuali.

Nel contesto specifico del progetto al quale ho preso parte, i partecipanti – selezionati tra studenti e giovani laureati in giurisprudenza, economia, sociologia, scienze politiche e ingegneria informatica – hanno lavorato in gruppi interdisciplinari, affiancati da esperti e facilitatori provenienti da tutta Europa. Il mio team era composto da giovani italiani, tedeschi e olandesi, con i quali ho avuto il privilegio di condividere un percorso intenso di confronto, co-progettazione e sperimentazione sul campo.

## **2. La metodologia adottata: il design thinking come strumento di inclusione**

Per affrontare in modo sistematico e creativo la complessa sfida dell'inclusione lavorativa, il gruppo di lavoro ha adottato la metodologia del *design thinking*, un approccio progettuale centrato sull'essere umano, che si è dimostrato particolarmente adatto a trasformare i bisogni latenti in soluzioni pratiche e replicabili. Si tratta di un metodo che unisce rigore ed immaginazione, struttura e intuizione, in un percorso iterativo finalizzato a ideare, testare e realizzare innovazioni guidate dall'ascolto profondo delle persone.

La forza di questa metodologia risiede nella sua capacità di mettere al centro la persona nella sua complessità, valorizzando l'esperienza vissuta, le emozioni, i limiti e le potenzialità di ciascuno. Nel nostro caso, il design thinking ha consentito di passare da un'analisi astratta del problema dell'esclusione lavorativa a una comprensione concreta e partecipata delle barriere che ostacolano l'accesso al lavoro, soprattutto per quei soggetti vulnerabili che vivono in condizioni di marginalità sociale, culturale e digitale. L'intero processo si è fondato sull'idea che l'intelligenza artificiale e gli strumenti digitali debbano essere messi al servizio dell'empowerment umano, e non viceversa.

Le cinque fasi classiche del design thinking – *empatizzare, definire, ideare, prototipare e testare* – sono state declinate secondo una logica di co-creazione continua, in dialogo costante con i beneficiari finali e gli attori locali. Ogni passaggio è stato vissuto non solo come un momento progettuale, ma anche come un'occasione di apprendimento reciproco, di crescita collettiva e di responsabilizzazione diffusa.

In questo spirito, il percorso progettuale si è articolato in una sequenza di passaggi distinti ma profondamente correlati, ognuno dei quali ha contribuito a raffinare progressivamente la comprensione del problema e a costruire risposte coerenti con i bisogni emersi. Le cinque fasi del design thinking non sono state applicate in modo meccanico, ma vissute come tappe dinamiche di un processo dialogico e trasformativo, dove ogni idea veniva costantemente messa alla prova dall'ascolto del contesto reale. Di seguito si offre una

ricostruzione dettagliata delle diverse fasi attraversate dal gruppo, evidenziandone le scelte metodologiche, le intuizioni progettuali e le implicazioni etiche.

## 1. Empatizzare: ascoltare prima di progettare

Ogni progetto autenticamente orientato al cambiamento sociale prende forma a partire da un atto di ascolto. È proprio dall'ascolto che nasce la possibilità di comprendere i bisogni più profondi, quelli che non si esprimono in numeri o indicatori, ma che si colgono nel tono della voce, nei silenzi tra le parole, negli sguardi abbassati e nei gesti esitanti. In questa consapevolezza ha preso avvio la fase dell'“empatizzare”, primo passaggio del design thinking, che ha rappresentato il cuore umano dell'intero progetto. Non si è trattato semplicemente di raccogliere dati o informazioni, ma di avvicinarsi con rispetto e apertura alla realtà vissuta dalle persone escluse dal mercato del lavoro nella città di Tilburg.

Attraverso interviste, osservazioni partecipanti, focus group e workshop esplorativi, il gruppo di lavoro ha costruito uno spazio sicuro in cui poter incontrare storie spesso dimenticate. Non erano semplici “utenti” o “beneficiari”: erano donne e uomini con alle spalle vite complesse, con sogni interrotti, con ferite invisibili e con una voglia sincera di sentirsi di nuovo parte di qualcosa. È emersa una realtà fatta di solitudini silenziose, percorsi spezzati e resilienze quotidiane.

Molti dei partecipanti hanno raccontato di sentirsi bloccati in un limbo esistenziale, sospesi tra un passato in cui avevano un'identità professionale e un presente in cui quella stessa identità sembra svanita. La perdita del lavoro non veniva percepita solo come un problema economico, ma come una ferita alla dignità personale. A questa si aggiungeva, in molti casi, una scarsa padronanza delle tecnologie digitali, un elemento oggi cruciale non solo per cercare lavoro, ma per accedere ai servizi, mantenere relazioni, partecipare alla vita pubblica. Per molti, anche scrivere un curriculum o aprire un account su un sito di job matching rappresentava un ostacolo insormontabile, che generava frustrazione, ansia, senso di inadeguatezza.

In mezzo a queste storie, il gruppo ha sentito l'urgenza di dare un volto concreto ai bisogni emersi. È nata così la figura di **Brigitte**, una *persona* immaginaria, ma costruita a partire da esperienze reali, e modellata come sintesi di molte voci ascoltate sul campo. Brigitte è

una donna di circa cinquant'anni, vive a Tilburg e ha lavorato per gran parte della sua vita nel settore manifatturiero. La sua storia è comune a molte altre: dopo la chiusura dello stabilimento dove era impiegata, non è più riuscita a reinserirsi nel mondo del lavoro. Non per mancanza di volontà o pigrizia, ma perché le competenze richieste sono cambiate, e il mondo intorno a lei si è trasformato troppo in fretta.

Brigitte non ha un forte supporto familiare, non possiede competenze digitali avanzate, si sente estranea ai meccanismi burocratici e spesso smarrita davanti alle logiche impersonali dei servizi pubblici. Ma Brigitte è anche altro: è una donna tenace, empatica, dotata di senso pratico e di spiccate capacità relazionali. È abituata a lavorare in gruppo, a risolvere problemi, ad affrontare imprevisti. Ha qualità umane preziose, ma che raramente vengono valorizzate nei processi automatizzati di selezione del personale.

È stata proprio questa dissonanza – tra il potenziale umano di Brigitte e la sua esclusione sistemica – a ispirare l'intero impianto progettuale. La sua storia è diventata **una bussola etica**, un riferimento concreto per mantenere al centro del lavoro non l'astrazione dell'"utente tipo", ma la ricchezza, le fragilità e le possibilità di una persona reale. Ogni decisione, ogni passaggio metodologico, ogni ipotesi di soluzione è stata costantemente confrontata con la domanda: "Cosa penserebbe Brigitte? In che modo questa scelta può davvero aiutarla?".

La sua figura ha permesso al gruppo di superare la tentazione di risposte generiche e di focalizzarsi su interventi mirati, sensibili, proporzionati. Sulla base dei dati forniti dall'amministrazione comunale, che indicavano un bisogno crescente di personale nel settore dell'assistenza agli anziani, è stato ipotizzato un percorso di riconversione professionale per Brigitte, fondato non su competenze tecniche avanzate, ma sulle sue qualità relazionali e sulla sua esperienza di cura informale. In un contesto spesso dominato da metriche e performance, Brigitte ha ricordato al gruppo che l'innovazione autentica non è mai disgiunta dall'umanità, e che un algoritmo può suggerire una strada, ma è l'ascolto a indicarci la direzione giusta.

Empatizzare, in questo progetto, ha significato **riconoscere la dignità di ogni storia, accogliere la complessità di ogni persona e costruire risposte che nascano dall'incontro, non dall'imposizione**. È da qui, da questo terreno fragile ma fecondo, che

ha preso forma il progetto: non da una lavagna bianca, ma da un volto, una voce, un nome. Da Brigitte.

## **2. Definire: dare un nome alla complessità, trasformare il problema in missione**

Dopo aver attraversato una fase di ascolto profondo e di immersione empatica nella realtà sociale di Tilburg, il gruppo di lavoro ha ritenuto fondamentale tradurre la complessità delle esperienze raccolte in una definizione chiara, condivisa e mobilitante del problema. Il compito non era semplice: le testimonianze raccolte durante interviste, workshop e osservazioni dirette restituivano un mosaico frammentato ma potente di disorientamento, invisibilità sociale, frustrazione e desiderio di riscatto. Era necessario ricomporre questo insieme in un'immagine coerente e operativa, che potesse guidare le successive fasi progettuali senza tradire la ricchezza e la densità umana del materiale raccolto.

È emerso con forza che una delle barriere principali all'inclusione lavorativa delle persone migranti e vulnerabili residenti a Tilburg risiedeva nell'incapacità di orientarsi nel sistema locale, non per mancanza di volontà o competenze, ma per l'assenza di strumenti accessibili e affidabili che potessero accompagnarle in questo percorso. Il problema, dunque, non era semplicemente tecnico o informativo: era strutturale, culturale, e in parte anche psicologico. Le persone incontrate dal gruppo non chiedevano soluzioni calate dall'alto, ma opportunità di accesso, riconoscimento, dialogo, accompagnamento.

La riflessione condivisa ha portato alla formulazione di un obiettivo progettuale ambizioso e allo stesso tempo concreto: sviluppare una piattaforma digitale capace di fungere da ponte tra l'individuo e il contesto territoriale, uno strumento intelligente ma umano, in grado di adattarsi ai bisogni reali e alle traiettorie personali degli utenti. Una piattaforma capace non solo di offrire indicazioni pratiche o suggerimenti formativi, ma soprattutto di valorizzare le competenze pregresse (anche se acquisite in contesti informali o difficilmente certificabili), di stimolare l'autoefficacia e la fiducia in sé stessi, di guidare la persona in un processo graduale di riappropriazione del proprio ruolo attivo nella società.

Da qui, la sfida progettuale si è cristallizzata attorno a una missione chiara: **trasformare l'esclusione in partecipazione, l'invisibilità in centralità, la vulnerabilità in potenziale umano**. Non si è trattato semplicemente di “definire un problema”, ma di dare forma condivisa a un’urgenza collettiva, trasformandola in una visione ispirante, capace di orientare con coerenza etica e pragmatismo ogni passaggio del lavoro. In questo senso, la fase di definizione ha rappresentato il cuore strategico del progetto: un punto di sintesi in cui l’ascolto si è fatto direzione, e la narrazione dei bisogni si è trasformata in architettura del cambiamento.

### **3. Ideare, dall’ascolto all’innovazione: co-progettare soluzioni digitali per l’inclusione**

Dopo aver definito con chiarezza la sfida progettuale, il gruppo è entrato nella fase dell’ideazione con uno spirito aperto, creativo e profondamente orientato al dialogo. Il momento era carico di energia: le esperienze raccolte e le narrazioni ascoltate nella fase empatica avevano lasciato un’impronta viva e urgente, e tutti i partecipanti sentivano la responsabilità, ma anche l’opportunità, di restituire dignità, voce e strumenti concreti alle persone che avevano ispirato il progetto. L’ideazione non è stata dunque solo un esercizio tecnico, ma un vero e proprio atto di restituzione e di cura.

In un clima di fiducia e ascolto reciproco, il gruppo ha incoraggiato la libera circolazione delle idee, stimolando ogni componente a proporre soluzioni anche audaci, fuori dagli schemi, capaci di sfidare l’esistente per immaginare scenari nuovi. La diversità dei background disciplinari – dal diritto alla sociologia, dall’economia all’ingegneria informatica – ha arricchito notevolmente il processo, permettendo di affrontare il problema da più angolazioni e di generare una pluralità di prospettive.

Tra le numerose proposte emerse, si sono distinti alcuni filoni ricorrenti: l’idea di un **chatbot multilingue** capace di offrire supporto in tempo reale a persone con competenze linguistiche limitate; lo sviluppo di **sistemi intelligenti di matching** tra competenze individuali e offerte formative o occupazionali locali; la creazione di **moduli di orientamento virtuale** e strumenti per l’**auto-valutazione delle soft e hard skills**, pensati per rendere l’utente consapevole delle proprie risorse e delle aree di

miglioramento. Non meno importante, è stata l'intuizione di costruire **interfacce semplici e inclusive** per la realizzazione di CV personalizzati, adattabili al contesto locale e validi anche in assenza di certificazioni ufficiali.

Dopo un intenso confronto interno e un costante dialogo con esperti del settore, stakeholder locali e facilitatori, il gruppo ha deciso di sintetizzare queste intuizioni in un'unica proposta coerente: lo sviluppo di una **piattaforma digitale integrata**, in grado di riunire tutte le funzionalità sopra elencate in un solo ecosistema user-friendly. La scelta non è stata solo strategica, ma anche profondamente valoriale: creare uno spazio digitale che non fosse solo funzionale, ma anche accogliente, intuitivo, modellabile sulle esigenze reali di ogni utente.

Questa piattaforma, immaginata come **un compagno di viaggio digitale**, non si limita a fornire strumenti tecnici, ma si propone di accompagnare l'utente lungo tutte le tappe del suo percorso di inclusione: dalla scoperta e valorizzazione delle proprie competenze, all'acquisizione di nuove abilità, fino al contatto diretto con il mondo del lavoro. La logica è quella dell'empowerment progressivo, costruito attraverso micro-passaggi, suggerimenti personalizzati, feedback continui e un'interfaccia empatica, pensata per ridurre l'ansia da prestazione e favorire la fiducia in sé stessi.

In questo modo, l'ideazione ha dato forma non solo a uno strumento tecnico, ma a un'**infrastruttura relazionale**, capace di restituire senso, direzione e possibilità a chi troppo spesso si è sentito invisibile o inadeguato. Ogni scelta tecnologica è stata guidata da un principio fondamentale: **mettere l'innovazione al servizio della persona, non il contrario.**

#### **4. Prototipare, dare forma all'idea: costruire uno strumento accessibile, umano, funzionale**

La fase di prototipazione ha rappresentato un momento di svolta nel percorso progettuale, in cui l'immaginazione si è fatta struttura e le intuizioni raccolte durante le fasi precedenti hanno trovato una prima traduzione concreta.

Il team ha lavorato con entusiasmo alla trasformazione delle idee in un prototipo digitale operativo, sviluppando un modello funzionale della piattaforma pensato per rispondere, in modo mirato e sostenibile, alle esigenze emerse nel confronto con i potenziali beneficiari.

Il prototipo è stato concepito come un ambiente digitale modulare, articolato in quattro aree fondamentali. La prima, dedicata alla **raccolta dei dati personali**, consentiva agli utenti di condividere informazioni rilevanti sul proprio percorso di vita, sulle esperienze lavorative passate e sulle preferenze individuali, garantendo sempre il rispetto della privacy. La seconda area era finalizzata alla **valutazione delle competenze**, sia formali sia informali, attraverso strumenti interattivi e questionari intuitivi, in grado di cogliere anche quelle soft skills spesso invisibili ai sistemi tradizionali. La terza sezione offriva una **generazione automatica di suggerimenti personalizzati**, grazie a un sistema intelligente che metteva in relazione il profilo dell'utente con opportunità formative concrete e realistiche. Infine, l'area dedicata all'**accesso al mercato del lavoro** proponeva annunci e offerte calibrate sulle competenze individuali, utilizzando un **algoritmo di matching** pensato per ridurre lo scarto tra desiderio e possibilità.

Particolare attenzione è stata riservata alla **progettazione dell'interfaccia grafica**, che doveva risultare accogliente, chiara e facilmente navigabile anche per persone con competenze digitali limitate. Il gruppo ha scelto un design semplice ma curato, caratterizzato da **icone intuitive, linguaggio visivo immediato e opzioni multilingue**, in modo da garantire una fruizione ampia, autonoma e rispettosa dei ritmi e delle capacità di ciascun utente. L'obiettivo non era solo quello di costruire uno strumento funzionale, ma anche uno **spazio digitale umanizzato**, dove ogni persona potesse sentirsi accolta, riconosciuta e valorizzata.

La prototipazione, lungi dall'essere una mera fase tecnica, ha rappresentato dunque un atto politico e relazionale: una prima concretizzazione del principio secondo cui la tecnologia, se ben orientata, può diventare alleata dell'inclusione, del riscatto sociale e della dignità umana.

## **5. Testare, il prototipo incontra la comunità: ascoltare, apprendere, migliorare**

La fase conclusiva del percorso progettuale ha rappresentato molto più di una semplice verifica tecnica: è stata l'occasione per ritornare idealmente alla comunità, chiudendo il cerchio avviato con l'ascolto empatico delle prime fasi. Dopo aver dato forma concreta alla piattaforma, il gruppo ha organizzato una serie di incontri pubblici e momenti di restituzione, durante i quali il prototipo è stato presentato ai principali attori che avevano supportato il progetto nelle sue fasi iniziali: dottorandi, operatori sociali, rappresentanti delle forze dell'ordine e altri referenti istituzionali del Comune di Tilburg.

Va precisato che la vera fase di implementazione e sperimentazione sul campo – con un coinvolgimento diretto e continuativo della cittadinanza – sarebbe potuta avvenire solo successivamente, qualora il progetto avesse ottenuto il finanziamento e avviato una collaborazione strutturata con l'amministrazione comunale. Questo traguardo è stato effettivamente raggiunto: il progetto è risultato vincitore e il gruppo sta attualmente mantenendo un rapporto attivo con l'amministrazione cittadina, con l'obiettivo di avviare concretamente la fase di applicazione e testing della piattaforma all'interno dei contesti reali per i quali è stata concepita.

In questa fase di gestazione, dunque, l'intento non è stato quello di testare il prodotto in senso stretto, bensì di raccogliere impressioni, valutazioni e suggerimenti da parte di interlocutori qualificati che avevano accompagnato o osservato con attenzione il processo di ideazione. Il confronto ha assunto la forma di interviste semi-strutturate, sessioni dimostrative interattive e test guidati, che hanno consentito di raccogliere feedback preziosi sia sul piano funzionale che valoriale. I partecipanti si sono mostrati coinvolti e propositivi, offrendo osservazioni puntuali che hanno evidenziato tanto i punti di forza della piattaforma quanto alcune aree di miglioramento, confermando l'efficacia dell'approccio partecipativo adottato lungo l'intero percorso.

Tra gli elementi più apprezzati vi sono stati **l'immediatezza del sistema**, la **chiarezza dell'interfaccia**, la **coerenza tra il profilo dell'utente e i suggerimenti ricevuti**, nonché la possibilità di **monitorare in autonomia i propri progressi** attraverso una dashboard personalizzata.

Parallelamente, sono emerse indicazioni significative per migliorare ulteriormente la piattaforma. In particolare, è stata rilevata l'esigenza di **rafforzare i meccanismi di trasparenza** nella gestione dei dati personali, e di **prevedere la presenza strutturale di figure di mediazione interculturale digitale**, capaci di accompagnare l'utente nel percorso e di facilitare l'interazione con lo strumento tecnologico, in particolare nei casi di bassa alfabetizzazione digitale o linguistica.

Questa fase non ha avuto solo una funzione correttiva. È stata anche un **atto di restituzione simbolica e politica**, che ha rafforzato il senso di **appartenenza collettiva al progetto**. L'esperienza di testing ha promosso una logica di **co-creazione**, trasformando il prototipo in **bene comune**, frutto di uno scambio continuo tra esperti e cittadini, tra tecnologia e vissuto. Attraverso il test, il progetto ha acquisito una nuova consapevolezza: quella di un'intelligenza artificiale che non pretende di guidare dall'alto, ma che **crece insieme alle persone**, adattandosi ai loro bisogni, rispettandone i ritmi e valorizzandone la dignità.

Attraverso le cinque fasi del design thinking, il gruppo è riuscito a trasformare i principi teorici dell'inclusione, dell'etica dell'innovazione e della centralità della persona in **un progetto vivo e tangibile**, capace di rispondere a bisogni reali e spesso invisibili. Ogni passaggio – dall'ascolto alla prototipazione, dalla definizione del problema al confronto con la comunità – ha rappresentato **un atto di cura e responsabilità**, volto a restituire dignità, voce e opportunità a chi, troppo spesso, rimane ai margini.

Il design thinking si è rivelato non solo uno strumento metodologico, ma **una postura etica**: ha saputo coniugare l'empatia con la creatività, l'immaginazione con il rigore, orientando la tecnologia verso finalità realmente umane. L'approccio human-centric ha evitato ogni forma di standardizzazione impersonale: l'obiettivo non era "inquadrali" o "classificarli", ma **accompagnare le persone in un percorso evolutivo**, riconoscendo e valorizzando la loro storia, le loro aspirazioni e i loro talenti.

La piattaforma concepita non è semplicemente un prodotto tecnologico, ma **uno spazio relazionale**: un ecosistema digitale che cresce con l'utente, lo ascolta, lo guida e si lascia trasformare dai suoi bisogni. In questa visione, l'intelligenza artificiale non è un fine, ma **un mezzo al servizio della crescita umana**, una tecnologia che si piega all'etica dell'inclusione e si nutre del dialogo con le vite reali.

L'esperienza vissuta a Tilburg mostra che la trasformazione digitale può essere anche **una trasformazione sociale**, a patto che venga costruita dal basso, con uno sguardo rispettoso sulla complessità delle persone e con la consapevolezza che l'innovazione vera nasce dall'incontro. In un'epoca in cui la tecnologia rischia spesso di disconnettere, questo progetto dimostra che si può e si deve usare il digitale per **ricucire legami**, restituire potere, e **costruire cittadinanza**

### **3. La piattaforma: struttura, funzionalità e finalità**

La piattaforma sviluppata all'interno del progetto Engage si configura come **un ecosistema digitale dinamico, inclusivo e personalizzabile**, pensato per rispondere alle reali esigenze delle persone migranti e vulnerabili residenti a Tilburg. Lontana da ogni logica astratta o standardizzata, essa nasce dall'ascolto diretto dei vissuti, dalle testimonianze emerse nella fase empatica e dalla volontà condivisa di costruire **uno strumento orientato all'empowerment individuale e comunitario**.

Fin dalla sua progettazione, la piattaforma ha integrato la convinzione che ciascun utente debba essere **ricosciuto come soggetto attivo**, portatore di esperienze, competenze e aspirazioni uniche. La sua funzione non è classificare, ma **accompagnare, valorizzare e facilitare** un cammino autonomo e consapevole verso l'inserimento lavorativo e formativo, rafforzando la fiducia in sé stessi e il senso di appartenenza alla comunità.

La struttura è articolata in **moduli interconnessi**, ognuno pensato per rispondere a una specifica dimensione del percorso di inclusione:

- **Profilazione iniziale e autovalutazione**: attraverso un'interfaccia semplice e intuitiva, l'utente viene guidato nella raccolta delle informazioni personali, nelle esperienze pregresse e nella valutazione delle proprie competenze, sia formali che

informali. Questo momento di riflessione attiva consente alla piattaforma di costruire un profilo completo, base su cui si fondano tutte le successive interazioni.

- **Raccomandazioni personalizzate:** la piattaforma utilizza un sistema intelligente di matching, alimentato da un algoritmo progettato in chiave etica e adattiva. Le raccomandazioni proposte non si limitano a suggerire offerte lavorative, ma includono corsi brevi, tirocini, attività di mentoring, eventi formativi e servizi di supporto (ad esempio legale, sanitario o psicologico), in stretta connessione con il contesto locale.
- **Interfaccia inclusiva e multilingue:** per rispondere alle esigenze di utenti con differenti livelli di alfabetizzazione digitale e linguistica, è stato adottato un design visivamente chiaro, supportato da icone esplicative, assistenti vocali, tutorial guidati e contenuti accessibili in più lingue, tra cui inglese, olandese e arabo.
- **Mappa delle opportunità locali:** uno degli strumenti più apprezzati del prototipo è il sistema cartografico interattivo che connette il profilo dell'utente a servizi e iniziative presenti nel territorio di Tilburg, suddivisi per ambiti e tipologie. Questa funzione rafforza l'incontro tra l'individuo e la rete sociale esistente, favorendo una reale integrazione nella vita civica della città.

La piattaforma non si limita a offrire servizi, ma si propone come **spazio relazionale trasformativo**, capace di crescere e migliorare nel tempo. In linea con i principi della co-progettazione, essa è concepita per evolvere grazie al contributo attivo degli utenti e alla collaborazione continua con stakeholder istituzionali e comunitari. Un **comitato etico di supervisione**, composto da rappresentanti locali ed esperti in tecnologie inclusive, vigila sul rispetto dei criteri di trasparenza, sicurezza e tracciabilità.

In questo quadro, **l'intelligenza artificiale non sostituisce l'umano, ma lo amplifica**, creando occasioni di incontro e orientamento, e restituendo centralità alla persona nella sua interezza. La piattaforma si fa così **ponte tra esigenze individuali e risposte collettive**, incarnando un nuovo paradigma di innovazione sociale, fondato sull'ascolto, la dignità e la responsabilità condivisa.

## 4. Verso una governance etica e partecipativa della piattaforma

Uno degli aspetti più delicati e insieme più significativi emersi nel corso del progetto *Engage* riguarda la governance della piattaforma digitale. Fin dalle prime fasi, il gruppo ha riconosciuto che una tecnologia concepita per accompagnare processi di inclusione non può limitarsi a fornire servizi: essa deve riflettere nei propri meccanismi di funzionamento gli stessi principi etici, partecipativi e relazionali che intende promuovere all'esterno.

Governare una piattaforma di questo tipo non significa semplicemente amministrare dati e funzionalità, ma costruire un'infrastruttura sociale in grado di custodire la fiducia degli utenti, garantire giustizia procedurale e restare permeabile alle trasformazioni della comunità. In quest'ottica, è emersa con forza la convinzione che l'inclusività non possa essere un obiettivo solo del sistema, ma debba costituirne la forma stessa.

A partire da queste premesse, il gruppo ha delineato un modello di governance profondamente innovativo, articolato attorno a quattro pilastri fondamentali:

- ***Trasparenza radicale***

Ogni utente ha il diritto di sapere come funziona la piattaforma che utilizza, soprattutto quando questa incide su aspetti così centrali come la ricerca di lavoro, la formazione o l'accesso ai servizi. Per questo, è stata proposta la creazione di un'infrastruttura informativa chiara, accessibile e multilingue, in grado di spiegare – anche attraverso strumenti visuali e interattivi – quali dati vengono raccolti, come sono trattati, quali criteri guidano le raccomandazioni dell'algoritmo e chi ne ha accesso. La trasparenza diventa così strumento di empowerment, e non mero adempimento tecnico.

- ***Partecipazione autentica***

Il gruppo ha sottolineato con forza la necessità di dare voce agli utenti non solo come destinatari, ma come co-protagonisti del sistema. Da qui nasce la proposta di costituire un User Board, un comitato misto composto da persone migranti, mediatori culturali, esperti di IA, operatori sociali e rappresentanti della società civile. Questo organismo avrebbe il compito di monitorare il funzionamento

della piattaforma, raccogliere esperienze d'uso, suggerire miglioramenti e contribuire a una valutazione periodica, offrendo una prospettiva critica e concreta dal basso.

- ***Supervisione etica indipendente***

Per prevenire derive discriminatorie o logiche opache, il progetto prevede l'istituzione di un comitato etico autonomo, composto da esperti di etica digitale, giurisprudenza, diritti umani, sociologia e tecnologia. Questo comitato avrebbe un ruolo centrale nella valutazione dell'impatto sociale dell'algoritmo, nella prevenzione dei bias sistemici e nella promozione dell'equità, redigendo un rapporto annuale sull'equità, la sicurezza e l'affidabilità del sistema. Si tratta di un presidio cruciale, soprattutto in un'epoca in cui i sistemi automatizzati rischiano di diventare opachi o autoreferenziali.

- ***Quadro normativo***

Ogni piattaforma che aspiri a essere giusta deve dotarsi di regole chiare, dinamiche e condivise. Per questo motivo è stato pensato uno "Statuto della piattaforma", un documento vivo, tradotto in più lingue, che stabilisca i principi fondativi del sistema, i diritti e doveri degli utenti, le responsabilità degli sviluppatori, le modalità di gestione e conservazione dei dati, e le procedure di reclamo e revisione. Questo statuto dovrebbe essere rivisto periodicamente, attraverso processi inclusivi, così da adattarsi nel tempo all'evoluzione dei bisogni e delle tecnologie.

Ma la riflessione sulla governance non si è fermata alla struttura interna della piattaforma. Il gruppo ha infatti riconosciuto che la sostenibilità di un progetto così ambizioso non può dipendere da un singolo attore, né essere affidata a finanziamenti estemporanei. Per questo motivo è stata proposta la creazione di una coalizione multi-stakeholder permanente, composta da enti pubblici, organizzazioni del terzo settore, università, imprese sociali e comunità locali. Questa rete dovrebbe garantire la stabilità economica e progettuale della piattaforma, con ruoli ben distinti ma profondamente integrati: dalla gestione tecnica alla valutazione di impatto, dalla manutenzione all'innovazione continua.

In definitiva, il modello proposto concepisce la governance come un processo continuo di co-creazione, e non come una semplice sorveglianza dall'alto. Non si tratta di “gestire” la piattaforma, ma di radicarla nella realtà sociale, culturale e giuridica del territorio, ascoltando i suoi cambiamenti, le sue fratture, le sue potenzialità. Così intesa, la governance diventa spazio di dialogo intersezionale, luogo di confronto tra saperi tecnici, esperienze di vita, competenze istituzionali ed esigenze collettive.

Come ribadito nel documento Engage, una piattaforma digitale basata sull'intelligenza artificiale non può dirsi riuscita solo perché funziona tecnicamente. Il suo successo va misurato nella capacità di generare fiducia, rafforzare il senso di appartenenza, garantire giustizia e dignità a chi la utilizza. E in questo senso, la governance diventa la sua anima più profonda: non un meccanismo di controllo, ma un patto di cura reciproca tra tecnologia e umanità.

## 5. Il gruppo contro al reskilling

Uno degli aspetti più originali e culturalmente rilevanti emersi nel progetto Engage riguarda il **rifiuto consapevole della logica del reskilling forzato**, ovvero di quella visione adattiva e performativa che riduce l'innovazione tecnologica a un imperativo di riconversione permanente del soggetto umano. In molte narrazioni mainstream, l'intelligenza artificiale è presentata come una forza neutra e inevitabile che ridefinisce costantemente i confini dell'occupabilità, imponendo ai lavoratori l'obbligo di “reinventarsi” secondo **standard dettati dalla produttività, dalla flessibilità e dalla compatibilità digitale**. In questo scenario, la persona viene implicitamente chiamata a correggere sé stessa, ad aggiornarsi incessantemente per non perdere contatto con un mondo del lavoro sempre più disincarnato, dove la velocità del cambiamento diventa la misura del valore.

Il gruppo ha invece compiuto una **scelta controcorrente e profondamente etica**: ha respinto la logica della riqualificazione intesa come “adeguamento coatto” e ha promosso un modello fondato sulla **valorizzazione della soggettività in tutta la sua complessità esistenziale**. Alla base del progetto vi è la convinzione che ogni persona possieda un patrimonio di esperienze, intuizioni, abilità relazionali e intelligenze non codificate

che **sfuggono ai modelli tradizionali di valutazione delle competenze**, spesso costruiti su griglie standardizzate, criteri astratti e metriche computazionali.

In questa prospettiva, l'intelligenza artificiale non è più uno strumento di selezione impersonale o di classificazione normativa, ma diventa **una tecnologia riflessiva**, capace di riconoscere le traiettorie individuali e di interagire con esse con rispetto, ascolto e sensibilità. Non è l'individuo a doversi piegare alle esigenze del sistema: è la tecnologia, al contrario, che deve farsi **strumento di emersione dell'umano**, di attivazione delle sue risorse latenti e di accompagnamento verso un'autentica autodeterminazione lavorativa e sociale.

Questo **ribaltamento di paradigma** ha orientato ogni fase del lavoro: dalla mappatura delle competenze informali alla progettazione di percorsi personalizzati, dalla scelta degli strumenti algoritmici alla definizione delle metriche valutative. Il gruppo Bridge ha operato in modo da rendere **visibili e valorizzabili anche quelle forme di sapere pratico, empatico, narrativo o situazionale** che non trovano riconoscimento nei sistemi canonici di certificazione o nel lessico del mercato. Le persone non sono state considerate soggetti da "riempire" di nuove abilità, ma **portatrici di un valore già presente**, che la tecnologia doveva semplicemente aiutare a rivelare.

In tal senso, l'intelligenza artificiale è stata trattata come **un'alleata epistemica ed etica**: uno strumento non per correggere ma per riconoscere, non per addestrare ma per interpretare, non per standardizzare ma per differenziare. Le soluzioni proposte non hanno mirato a colmare un "deficit" del soggetto rispetto a una norma ideale, bensì a **costruire un contesto in cui potesse emergere la pluralità delle intelligenze umane**, nella loro dimensione biografica, culturale e sociale.

Questo approccio ha avuto anche una chiara valenza politica: **rifiutare il reskilling forzato significa rifiutare un modello di società in cui l'individuo è valutato unicamente per la sua capacità di adattamento ai bisogni mutevoli del sistema economico**. Si è voluta invece restituire centralità alla libertà delle persone, creando uno spazio protetto in cui potessero **riscoprire e ridefinire il proprio valore al di fuori delle logiche produttivistiche**. La competenza, in questo quadro, non è più un parametro da aggiornare continuamente per restare competitivi, ma una **qualità dinamica e**

**relazionale**, che si sviluppa nel dialogo con la comunità, nella costruzione di senso e nel riconoscimento reciproco.

In definitiva, il progetto pilota ha dimostrato che **una tecnologia intelligente può essere anche profondamente umana**, se guidata da valori chiari, da un'attenzione autentica alla dignità del soggetto e da una visione della formazione non come normalizzazione, ma come occasione di fioritura personale. Il rifiuto del reskilling forzato si è così trasformato in **una proposta culturale potente e necessaria**, capace di restituire al lavoro la sua dimensione di senso e all'intelligenza artificiale la sua vocazione etica più alta.

## **6. Implicazioni etico-sociali del progetto Engage**

Il progetto Engage non si è limitato a sviluppare un'infrastruttura tecnologica avanzata per migliorare l'accesso ai servizi da parte di migranti e residenti locali. Ha saputo spingersi oltre, affrontando con consapevolezza **questioni etiche, sociali e politiche di grande profondità**, nella convinzione che ogni innovazione digitale, se non accompagnata da un pensiero critico, rischi di rafforzare esclusioni anziché abatterle. L'inclusione digitale, in questa prospettiva, non è mai un mero fatto tecnico o logistico: è un processo complesso di cittadinanza attiva, **di riconoscimento dell'altro come soggetto di diritti e desideri**, e di trasformazione delle relazioni tra individui, istituzioni e tecnologie

Fin dalle prime fasi di ideazione, il gruppo di lavoro ha individuato nel rischio di **disuguaglianze algoritmiche** una delle criticità più insidiose. Un sistema di intelligenza artificiale, se addestrato su dati parziali, distorti o culturalmente omogenei, può generare effetti opposti a quelli desiderati: invece di facilitare l'accesso, può produrre esclusione, invece di valorizzare la diversità, può rafforzare stereotipi. Profilazioni automatizzate basate su modelli standardizzati rischiano di penalizzare chi ha un vissuto non lineare o non conforme agli schemi formativi e lavorativi convenzionali, come accade spesso nei percorsi migratori. Per questo il gruppo ha posto al centro del lavoro **un modello inclusivo di data management**, che prevede la revisione continua delle fonti, la partecipazione attiva degli utenti alla costruzione dei propri profili e il controllo trasparente degli output generati dal sistema.

A questa riflessione si è affiancata una profonda attenzione al tema della **fiducia nelle tecnologie pubbliche**. Per molte persone migranti, il rapporto con i sistemi digitali istituzionali è segnato da timori concreti: paura della sorveglianza, dell'utilizzo improprio dei dati, della perdita di controllo sulla propria identità digitale. In risposta a questa diffidenza strutturale, il progetto ha costruito un modello di **empowerment digitale** fondato su trasparenza radicale, accesso all'informazione e capacità di scelta. Ogni utente deve sapere, in modo chiaro e comprensibile, cosa accade ai propri dati, come vengono usati, e quali sono le implicazioni delle decisioni algoritmiche che lo riguardano. **La trasparenza non è stata concepita come adempimento formale**, ma come condizione essenziale per una partecipazione piena, consapevole e libera.

La barriera della lingua, dell'alfabetizzazione digitale e dei codici culturali rappresenta spesso il primo e più potente ostacolo all'inclusione. Per questo la piattaforma è stata progettata in modalità multilingue, con interfacce semplificate, moduli audiovisivi e tutorial interattivi. Ma il gruppo è andato oltre: ha riconosciuto che **nessuna tecnologia, per quanto sofisticata, può sostituire l'interazione umana**. Da qui la decisione di affiancare alla dimensione digitale la presenza attiva di **facilitatori interculturali e psicologi del benessere**, figure capaci di interpretare i bisogni impliciti, di accompagnare gli utenti in un percorso di senso, e di favorire processi di apprendimento reciproco tra migranti e operatori locali.

In questa visione inclusiva, **le comunità locali non sono semplici spettatrici del cambiamento**, ma protagoniste di un processo trasformativo che le coinvolge in prima persona. L'inclusione, infatti, non è mai un processo unidirezionale: non consiste nel "portare dentro" i migranti nei sistemi esistenti, ma nel **trasformare quei sistemi a partire dal riconoscimento delle alterità che li attraversano**. Engage ha quindi previsto funzionalità specifiche per attivare il tessuto sociale locale: spazi per il volontariato, strumenti per la co-formazione tra pari, eventi interculturali e ambienti di scambio di buone pratiche. Si è trattato, in altre parole, di **favorire un dialogo costante tra vecchi e nuovi cittadini**, in cui la tecnologia diventasse occasione di incontro, e non barriera.

Infine, la riflessione si è spinta verso **una dimensione antropologica e filosofica**, interrogando il ruolo stesso dell'intelligenza artificiale nella costruzione dell'umano.

L'IA non è una macchina neutra: ogni scelta progettuale porta con sé **valori impliciti, visioni del mondo, priorità culturali**. Il nostro gruppo ha scelto di non eludere queste domande, ma di affrontarle frontalmente, proponendo un'idea di intelligenza artificiale come **tecnologia relazionale**, chiamata non solo a funzionare bene, ma a **rispecchiare un ideale di giustizia sociale, di equità e di riconoscimento della dignità**. In questo senso, l'“inclusione algoritmica” non è stata trattata come un obiettivo tecnico, ma come **una sfida politica e culturale**, che impone di ripensare il rapporto tra automazione e libertà, tra efficienza e umanità, tra innovazione e giustizia.

## **7. Conclusione: Un modello replicabile di innovazione sociale e giustizia digitale**

Nel delineare e realizzare il progetto Engage, abbiamo scelto con convinzione di abbracciare **una visione antropocentrica dell'innovazione**, che pone la persona – con la sua dignità, complessità e unicità – al centro dei processi tecnologici, sociali e normativi. Questa impostazione non è frutto di un'intuizione estemporanea, ma si radica in una visione matura e responsabile del futuro, in linea con una riflessione etica e filosofica oggi sempre più centrale nel dibattito internazionale. Figure autorevoli come **Luciano Floridi, Virginia Dignum e Laura Palazzani** hanno più volte richiamato l'urgenza di un'intelligenza artificiale concepita *per e con l'uomo*, e non in sua sostituzione o estraneazione. Engage si inserisce esattamente in questo solco, trasformando quella che spesso rimane una dichiarazione di principio in un'esperienza progettuale concreta.

Il progetto realizzato a Tilburg rappresenta, in questo senso, **un esempio emblematico di co-progettazione avanzata**, in cui istituzioni, università, comunità migranti e società civile si sono incontrate non solo per condividere risorse, ma per **generare visioni comuni**, fondate sull'ascolto, la corresponsabilità e il rispetto reciproco. La piattaforma elaborata non è un mero prodotto tecnologico, bensì **un ecosistema relazionale**, un ambiente vivo e dinamico che incarna un'idea profondamente umana di innovazione.

L'approccio human-centric, il ricorso rigoroso e creativo al design thinking, la strutturazione di una governance partecipativa e multilivello, l'integrazione di riflessioni

giuridiche, etiche e antropologiche: tutto ciò ha contribuito a **delineare un modello avanzato e replicabile di giustizia digitale inclusiva**, capace di affrontare le sfide della contemporaneità con strumenti nuovi, ma radicati in valori antichi. L'esperienza di Engage dimostra che la tecnologia può essere al servizio dell'umano, ma solo se guidata da una tensione etica costante, da una progettazione orientata al bene comune e da un'infrastruttura relazionale che valorizzi la partecipazione attiva dei soggetti coinvolti.

I risultati ottenuti dal gruppo Engage indicano una via possibile: **una terza via tra tecnocentrismo e immobilismo**, capace di tenere insieme l'innovazione digitale e la coesione sociale, lo sviluppo delle IA e la tutela dei diritti fondamentali. L'esperienza condotta a Tilburg non è un caso isolato, ma **un paradigma generativo**, che può ispirare politiche pubbliche, progetti territoriali e processi normativi in tutta Europa. In un'epoca segnata dalla corsa alla digitalizzazione e dalla moltiplicazione delle vulnerabilità, Engage mostra che **“fare tecnologia” non significa abdicare alla complessità**, ma abitarla con cura, immaginazione e responsabilità.

Nel quadro complessivo di questa tesi, il caso Engage si configura come **la concretizzazione dei principi analizzati nei capitoli precedenti**: dalla trasparenza alla spiegabilità, dalla non discriminazione alla responsabilità, dalla dignità alla partecipazione attiva. Non si tratta di una semplice applicazione sperimentale o di un esercizio progettuale isolato, ma di **una forma di etica incarnata**, che prende corpo in uno strumento digitale reale e funzionale. Engage rappresenta la prova che è possibile trasformare visioni teoriche in pratiche trasformative, e che l'innovazione non è neutra, ma può essere **fatta bene, fatta con e fatta per**.

Ciò che emerge da questa esperienza è un insegnamento profondo: l'innovazione digitale **non è automaticamente emancipatoria**, ma può diventarlo se nasce da un'autentica attenzione ai bisogni delle persone, se riconosce il valore delle relazioni umane, se accetta di essere progettata in modo consapevole e condiviso. Engage ci ricorda che non dobbiamo limitarci ad adattarci alle tecnologie emergenti, ma dobbiamo **rivendicare il diritto e il dovere di disegnarle in base alla società che desideriamo costruire**. In questa prospettiva, il coinvolgimento delle comunità locali, l'adozione di metodologie partecipative, la valorizzazione delle competenze trasversali e

la promozione dell'empowerment individuale non sono aspetti accessori, ma **componenti strutturali di un nuovo modo di intendere la giustizia nell'era digitale.**

La piattaforma Engage, dunque, non è solo un servizio innovativo. È **un simbolo potente di ciò che può nascere dall'incontro tra intelligenza artificiale e intelligenza sociale;** tra algoritmi e biografie; tra dati e diritti. È un ponte tra il mondo che abbiamo e il mondo che vogliamo. E in questo senso, Engage è molto più di un progetto: potrebbe diventare **manifesto per una nuova stagione dell'innovazione europea,** in cui la tecnologia diventa finalmente uno strumento di umanità, solidarietà e futuro condiviso.

L'esperienza di Tilburg trova un riscontro diretto e significativo nelle riflessioni contenute nell'**Opinione n. 30 dell'European Group on Ethics in Science and New Technologies (EGE), *Future of Work, Future of Society* (2018),** che rappresenta uno dei contributi più autorevoli sul rapporto tra tecnologia, lavoro e giustizia sociale. In questo documento si evidenzia come l'automazione, l'intelligenza artificiale e la digitalizzazione stiano trasformando radicalmente il ruolo del lavoro, con effetti ambivalenti: da un lato, possibilità di alleggerire mansioni faticose o pericolose e di ampliare la produttività; dall'altro, rischi di precarizzazione, polarizzazione e perdita di dignità lavorativa. L'EGE sottolinea che il nodo non è tanto "se" la tecnologia trasformerà il lavoro, quanto "come" le società europee decideranno di governare questo processo, affinché il futuro del lavoro sia anche il futuro di una società più giusta.

Un aspetto centrale dell'Opinione è la necessità superare una logica di **upskilling individuale** per approdare ad una strategia più ampia di "**societal upskilling**". Non si tratta solo di chiedere ai singoli lavoratori di aggiornare di aggiornare incessantemente le proprie competenze, ma di costruire collettivamente quadri normativi capaci di garantire a tutti condizioni di vita dignitose, indipendentemente dalla posizione occupazionale. In assenza di questo impegno condiviso, l'"aggiornamento continuo" rischia di diventare un meccanismo punitivo verso chi non riesce a tenere il passo con le richieste di competitività. È esattamente questo il paradigma che il progetto *Engage* ha voluto sovvertire, rifiutando la logica del reskilling coatto e proponendo invece una

valorizzazione delle competenze informali, delle capacità relazionali e delle esperienze biografiche non codificate.

Il documento dell'EGE pone inoltre l'accento sulla **non strumentalità del lavoro**: non basta che esso produca reddito, ma deve rappresentare anche uno spazio di espressione, riconoscimento e appartenenza sociale. In quest'ottica, attività non retribuite come la cura familiare o il volontariato assumono un valore pari a quello del lavoro salariato, in quanto contribuiscono alla coesione e al benessere collettivo. La piattaforma sviluppata a Tilburg rispecchia questa impostazione, riconoscendo e rendendo visibili forme di capitale umano e sociale spesso invisibili ai meccanismi tradizionali di selezione.

Un ulteriore punto di contatto riguarda la gestione dei **dati e della privacy**. L'EGE denuncia con forza i rischi di una crescente "datafication" del lavoro, che può tradursi in forme di sorveglianza indebita e di valutazione algoritmica standardizzata, lesive della dignità e dell'autonomia dei lavoratori. Engage, nel proporre un modello di governance etica e partecipativa, ha cercato di rispondere a questa sfida prevedendo trasparenza radicale, comitati etici indipendenti e la presenza attiva di mediatori interculturali, così da prevenire discriminazioni e rafforzare la fiducia degli utenti.

Infine, il documento delinea una visione del futuro del lavoro come "**questione morale**" prima ancora che economica, richiamando valori quali dignità, giustizia, uguaglianza, solidarietà e autonomia come criteri guida. Proprio questi valori hanno costituito la bussola dell'esperienza di Tilburg, che non ha considerato l'intelligenza artificiale un fine in sé, ma uno strumento per amplificare le capacità umane e promuovere inclusione. In questo senso, *Engage* si configura come una traduzione pratica delle linee etiche europee: una **prova di concetto** che mostra come sia possibile piegare la tecnologia a fini di giustizia sociale, facendo dell'innovazione non un fattore di esclusione, ma un'occasione di cittadinanza attiva e di coesione.<sup>333</sup>

---

<sup>333</sup> European Group on Ethics in Science and New Technologies (EGE), *Future of Work, Future of Society. Opinion No. 30*, Publications Office of the European Union, Luxembourg, 2019.

## CONCLUSIONI

L'intelligenza artificiale si presenta come una delle sfide più radicali del nostro tempo. Non si tratta di una semplice innovazione tecnica, ma di un fenomeno capace di ridefinire i paradigmi sociali, culturali e giuridici entro i quali si colloca la vita umana e lavorativa. Essa sollecita il diritto e l'etica a interrogarsi non solo sulla regolazione dei rischi, ma soprattutto sulle condizioni che rendono possibile una convivenza realmente giusta, inclusiva e rispettosa della dignità delle persone.

Il lavoro è uno degli ambiti in cui l'impatto dell'IA risulta più evidente. Non tanto per il numero di posti che potrebbero essere persi o creati, quanto per la trasformazione qualitativa che l'IA impone ai contenuti professionali, alle dinamiche organizzative e ai rapporti di potere all'interno delle imprese. Le nuove tecnologie possono liberare i lavoratori da mansioni ripetitive e gravose, ma possono anche generare nuove forme di sorveglianza, di discriminazione algoritmica e di precarizzazione. La questione centrale, dunque, non è se l'IA sostituirà o meno il lavoro umano, bensì quale tipo di lavoro umano sopravviverà, con quali garanzie di tutela e con quale riconoscimento di valore.

Da questo punto di vista, il diritto si trova nella posizione di dover accompagnare e governare un cambiamento che rischia altrimenti di essere guidato unicamente da logiche di mercato ed efficienza. L'AI Act rappresenta il tentativo europeo di dare una risposta organica a queste sfide, trasformando principi etici come trasparenza, spiegabilità, non discriminazione e controllo umano in vincoli giuridici. Tuttavia, nessuna normativa, per quanto sofisticata, può esaurire la complessità del rapporto tra uomo e tecnologia: ciò che occorre è un orizzonte culturale e valoriale che faccia da cornice alle regole.

In questa direzione, il richiamo alla vulnerabilità e alla fiducia offre una chiave interpretativa preziosa. Come osserva Macioce, la vulnerabilità non è un difetto da correggere né una condizione di mera debolezza, ma il tratto costitutivo dell'umano, ciò che rende possibile l'apertura all'altro e la costruzione di legami sociali fondati sulla fiducia reciproca.<sup>334</sup> Proprio l'IA, che ha il potere di amplificare o ridurre le possibilità di vita delle persone, diventa banco di prova della capacità collettiva di custodire questa

---

<sup>334</sup> MACIOCE, C., *Vulnerabilità e fiducia*, in «Munera», n. 2/2022, p. XX.

vulnerabilità. Se progettata e regolata in modo inclusivo, può divenire strumento di rafforzamento della fiducia sociale, capace di promuovere coesione, solidarietà e nuove forme di partecipazione. Se lasciata invece alle sole logiche del profitto e della massimizzazione dell'efficienza, rischia di accrescere insicurezze, emarginazioni e nuove forme di disuguaglianza.

L'esperienza del progetto *Engage*, sviluppato a Tilburg, rappresenta un esempio concreto di come l'intelligenza artificiale possa essere orientata a fini di inclusione sociale e lavorativa. La piattaforma ideata ha dimostrato che è possibile utilizzare l'IA non per accentuare la distanza tra i soggetti, ma per favorire l'incontro tra domanda e offerta di lavoro, promuovendo opportunità anche per le categorie più vulnerabili. In questo senso, *Engage* incarna l'idea che la tecnologia, se guidata da principi etici e da un disegno giuridico adeguato, possa contribuire a rafforzare il tessuto fiduciario della società e a ridurre le disuguaglianze invece di moltiplicarle.

Le conclusioni che emergono sono dunque chiare: l'etica e il diritto dell'intelligenza artificiale non possono limitarsi alla neutralità tecnica, né esaurirsi in una mera gestione del rischio. Essi devono assumere la vulnerabilità come criterio guida per orientare l'innovazione e la fiducia come fine da perseguire. Solo così sarà possibile costruire un modello di cittadinanza digitale che non si fondi sulla selezione e sull'esclusione, ma sull'inclusione e sulla giustizia sociale. Solo così l'IA potrà diventare non una minaccia alla dignità umana, ma una delle vie attraverso cui rafforzarla e rilanciarla, trasformando il progresso tecnologico in progresso umano.

## BIBLIOGRAFIA

ADAMS-PRASSL, J., *What if your boss was an algorithm? Economic incentives, legal challenges, and the rise of artificial intelligence at work*, in *Comparative Labor Law & Policy Journal*, vol. 41, 2019.

AA. VV., *Il futuro del lavoro si chiama “smart working”? Riflessioni e prospettive*, in «Rivista trimestrale di Organizzazione Aziendale».

ALLEGRI, M. R., *L’ambiguo principio (anche costituzionale?) della trasparenza algoritmica fra tecnologia, diritto e linguaggio*, in *Saggi*, 2021.

BIASI, M., *La presunzione relativa di subordinazione nella “Direttiva Piattaforme”*: *brevi spunti*, in *Lavoro Diritti Europa*, n. 4, 2024.

CALIGIURI, A. (a cura di), *Responsabilità e AI*, Giappichelli, Torino, 2022.

CASIRAGHI, S., VAN DIJK, N., *Ethics reviews in the European Union. Implications for the governance of scientific research in times of data science and Artificial Intelligence*, in *Law, Innovation and Technology*, vol. 16, n. 1, 2024.

CIPRIANI, A., GRAMOLATI, A., MARI, G. (a cura di), *Il lavoro 4.0. La Quarta Rivoluzione Industriale e le trasformazioni delle attività lavorative*, Firenze, Firenze University Press, 2018.

CIRONE, E., *L’AI Act e l’obiettivo (mancato?) di promuovere uno standard globale per la tutela dei diritti fondamentali*, in «Quaderni AISDUE – Rivista quadrimestrale», fasc. spec. n. 2/2024.

COLLECCHIA, G., *Apriamo la scatola nera: l’intelligenza artificiale spiegabile*, in *Recenti Progressi in Medicina*, n. 11, 2021.

CONTALDI, G., *Intelligenza artificiale e dati personali*, in «Ordine internazionale e diritti umani», 2021.

COSTA, M., *Il lavoro nella transizione digitale e robotica: le dimensioni della capability to choice e quella to voice per la libertà realizzativa e di scelta del lavoratore*, in *MeTis. Mondi educativi. Temi, indagini, suggestioni*, vol. 14, n. 1, 2024.

COVELLI, A., *Privacy e lavoro: la gestione algoritmica dei rapporti lavorativi*, in COVELLI, R., *Lavoro e intelligenza artificiale: dai principi di trasparenza algoritmica al diritto alla conoscibilità*, in *Labour & Law Issues*, vol. 9, n. 1, 2023.

DEL GIGLIO, I., *Valutazione della performance mediante tecniche di People Analytics. Privacy in employment, controllo o innovazione*, in «Journal of Ethics and Legal Technologies», vol. 3, n. 2, novembre 2021.

- DESCARTES, R., *Discorso sul metodo*, parte IV, 1637; ed. it. a cura di E. Garin, Laterza, Bari, 2009.
- DEVON, R., VAN DE POEL, I., *Design Ethics: The Social Ethics Paradigm*, in *International Journal of Engineering Education*, vol. 20, n. 3, 2004.
- DI TANO, F., *Intelligenza Artificiale, ricerca scientifica e adeguamento etico-giuridico*, in *i-lex – Rivista di Scienze Giuridiche, Scienze Cognitive ed Intelligenza Artificiale*, vol. 18, n. 1, 2025.
- DERIU, P., RACIOPPI, S., *Riflessioni in tema di intelligenza artificiale e attività di vigilanza*, in «Quaderni FinTech», n. 15, CONSOB, agosto 2025.
- DONINI, A. M., COVELLI, R., *Lavoro e intelligenza artificiale: dalla trasparenza alla conoscibilità*, in *Labour & Law Issues*, vol. 9, n. 1, 2023.
- EUROPEAN GROUP ON ETHICS IN SCIENCE AND NEW TECHNOLOGIES (EGE), *Future of Work, Future of Society. Opinion No. 30*, Publications Office of the European Union, Luxembourg, 2019.
- FLORIDI, L., et al., *AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*, in «Minds and Machines», vol. 28, 2018.
- FONDAZIONE ASTRID, *Quaderno n. 17 – L’AI Act: verso una regolazione europea dell’intelligenza artificiale*, estratto.
- FONDAZIONE BASSETTI, *AI e Sanità. Tecnologie, etica, responsabilità*, in LIMONE, D. A. (a cura di), *AI & Human Rights*, Roma, 2023.
- GARANTE PER LA PROTEZIONE DEI DATI PERSONALI, *Linee guida sui trattamenti di dati personali per finalità di gestione del rapporto di lavoro in ambito pubblico e privato*, Roma, 5 giugno 2019.
- GAUDIO, G., *Le discriminazioni algoritmiche*, in *Lavoro Diritti Europa*, n. 1, 2024.
- GENTILINI, V., FILOSA, R., *La tutela della salute e sicurezza del lavoro nello smart working*, Working Paper n. 20.
- GRAMOLATI, A., CIPRIANI, A., MARI, G. (a cura di), *Il lavoro 4.0. La Quarta Rivoluzione industriale e le trasformazioni delle attività lavorative*, Firenze University Press, Firenze, 2018.
- GRASSI, E., *Etica e intelligenza artificiale*, Roma, Armando Editore, 2021.
- GREEN, A., *Artificial Intelligence and the Changing Demand for Skills in the Labour Market*, OECD Artificial Intelligence Papers, n. 14, OECD Publishing, Paris, 2024.

- HEILMAN, M. E., CALEO, S., MANZI, F., *Women at Work: Pathways from Gender Stereotypes to Gender Bias and Discrimination*, in *Annual Review of Organizational Psychology and Organizational Behavior*, vol. 11, 2024.
- HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A Definition of AI: Main Capabilities and Scientific Disciplines*, European Commission, Bruxelles, 8 aprile 2019.
- KRAUS, C., GIACONI, C., D'ANGELO, A., *Different bodies: ibridazioni e trasformazioni nelle storie biografiche di persone con disabilità*, in *MeTis. Mondi educativi. Temi, indagini, suggestioni*, vol. 12, n. 2, 2022.
- KRAUS, C., GIACONI, C., D'ANGELO, A., *Abitare le differenze: il digital divide nelle disabilità*, in *MeTis*, vol. 12, n. 2, 2022.
- LAUANDE RODRIGUES, P., *Trabalho em plataformas digitais, subordinação algorítmica e presunção do vínculo empregatício*, in *Lavoro e Previdenza Oggi*, n. 9–10, 2022.
- LONGATO, E., *Bias e fairness nelle applicazioni cliniche dell'intelligenza artificiale*, Tesi di laurea in Ingegneria Biomedica, Università di Padova, a.a. 2023–2024.
- MACIOCE, C., *Vulnerabilità e fiducia*, in «Munera», n. 2/2022.
- MACERATINI, A., *Dati digitali e intelligenza artificiale. Sfide etiche e giuridiche*, in «Tigor. Rivista di scienze della comunicazione e di argomentazione giuridica», a. XIV, n. 2, 2022.
- MARTONI, M., *Spiegare le pieghe. Alcune osservazioni su trasparenza e spiegabilità*, in «L'Ircocervo», n. 1, 2024.
- MARAZZA, M., *Dei poteri (del datore di lavoro), dei controlli (a distanza) e del trattamento dei dati (del lavoratore)*, in *Biblioteca 20 Maggio*, n. 1, 2016.
- MOORE, P. V., UPCHURCH, M., WHITTAKER, X. (a cura di), *Humans and machines at work*, Palgrave Macmillan, Cham, 2018.
- OECD, *Policy Observatory: Implementing the OECD AI Principles*, Paris, 2021.
- OECD, *Recommendation of the Council on Artificial Intelligence*, OECD Legal Instruments, OECD/LEGAL/0449, 2019.
- PALAZZANI, L., *Etica della regolazione dell'AI*, in *Etica per le tecnologie emergenti*, CNR Edizioni, Roma, 2023.
- PICCINI, M., CARPIGNANO, A., *I fattori umani nel progetto di sistemi di controllo e di interfacce uomo-macchina in sistemi complessi altamente automatizzati*, Politecnico di Torino – Dipartimento di Energetica.

- PIETROPAOLI, S., *Dalla sorveglianza al controllo: la parabola della governamentalità algoritmica*, in «Rivista Italiana di Informatica e Diritto», n. 2, 2024.
- PIRO, P., *Intelligenza artificiale (IA) e mercati del lavoro*, Preprint, ResearchGate, maggio 2025.
- PIZZETTI, F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli, Torino, 2021.
- RANDSTAD, *Workmonitor Pulse – Speciale Intelligenza Artificiale e Inclusione*, 2024.
- RESSEGUIER, A., UFERT, F., *AI research ethics is in its infancy: the EU's AI Act can make it a grown-up*, in *Research Ethics*, vol. 20, n. 2, 2024.
- RICCI, A., *Il diritto alla disconnessione tra benessere del lavoratore e nuove forme di subordinazione digitale*, in *Rivista italiana di diritto del lavoro*, 2022.
- RUFFOLO, U. (a cura di), *Intelligenza artificiale e responsabilità. Atti del Convegno*, Università per Stranieri di Perugia, 29 novembre 2017, Giappichelli, Torino, 2018.
- SARRA, C., *L'uso di dati biometrici nelle procedure di reclutamento al lavoro mediante strumenti di Intelligenza Artificiale. Difficoltà normative multilivello*, in «Journal of Ethics and Legal Technologies», vol. 4, n. 2, novembre 2022.
- SCHWAB, K., *La quarta rivoluzione industriale*, Milano, Franco Angeli, 2017.
- SEGHEZZI, F., *Tecnologie e lavoro: l'Italia e la quarta rivoluzione industriale*, in *Quaderni di Adapt*, 2020.
- SOMALVICO, M., *L'Intelligenza Artificiale*, Rusconi Editore, Milano, 1987.
- TOMMASI, S., *Algoritmi e discriminazione: la responsabilità dell'umano nella macchina*, in ALTROCONSUMO (a cura di), *AI e Discriminazione*, Altroconsumo Edizioni, Milano, 2020.
- THALION, M., *The right to disconnect: work-life balance and digital labour law in Europe*, in «International Journal of Comparative Labour Law», 2021.
- TRENTHMAN, S., LARWOOD, L., *Gender Discrimination and the Workplace: An Examination of Rational Bias Theory*, in *Sex Roles*, vol. 38, nn. 1-2, 1998.
- UNESCO, *ChatGPT and Artificial Intelligence in Higher Education – Quick Guide*, Paris, 2023.
- UNESCO, *Generative AI and the Future of Education*, a cura di S. Giannini, Paris, 2023.
- UNESCO, *Guidance for Generative AI in Education and Research*, Paris, 2023.

UNESCO, *Moving from digital divide to digital inclusion: key principles for digital inclusion*, Paris, 2022.

UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, Paris, 2021.

ZAPPALÀ, L., *Informatizzazione dei processi decisionali e diritto del lavoro: algoritmi, poteri datoriali e responsabilità del prestatore nell'era dell'intelligenza artificiale*, in «Biblioteca "20 Maggio"», n. 2, 2021.

## SITOGRAFIA

AGENDA DIGITALE, *Un aiuto dalla tecnologia per muoversi in città: Maps for Easy Paths (MEP)*, disponibile su: <https://www.agendadigitale.eu>, ultimo accesso: 1 settembre 2025.

ANALYTICS INSIGHT, *The Global Artificial Intelligence Trends 2020*, cit. in I-COM, *L'intelligenza artificiale nei diversi settori: dove siamo e dove stiamo andando*, 2020, disponibile su: <https://www.i-com.it/2020/11/26/intelligenza-artificiale-settori/>, ultimo accesso: 1 settembre 2025.

COMMISSIONE EUROPEA, *Cosa si intende per intelligenza artificiale (IA)*, in “Digital Strategy”, disponibile su: <https://digital-strategy.ec.europa.eu/it/policies/artificial-intelligence>, ultimo accesso: luglio 2025.

IBM, *Cos'è il deep learning?*, disponibile su: <https://www.ibm.com/it-it/topics/deep-learning>, ultimo accesso: luglio 2025.

OSSERVATORIO BIG DATA & BUSINESS ANALYTICS, *Big Data e Business Analytics*, Politecnico di Milano, disponibile su: <https://www.osservatori.net/it/ricerche/osservatori-attivi/big-data-business-analytics>, ultimo accesso: luglio 2025.

REUTERS, *Amazon scraps secret AI recruiting tool that showed bias against women*, 11 ottobre 2018, disponibile su: <https://www.reuters.com>, ultimo accesso: 1 settembre 2025.

TJOA, M., *A Step by Step Backpropagation Example*, in “Towards Data Science”, disponibile su: <https://towardsdatascience.com/a-step-by-step-backpropagation-example-915fcf6cbee6>, ultimo accesso: luglio 2025.

