



La legge sulla IA tra finalità, prevenzione e prova nella tutela della salute e sicurezza sul lavoro

MICHELE SQUEGLIA

Università degli Studi di Milano Statale

Associato di Diritto del lavoro, previdenza sociale e sicurezza

michele.squeglia@unimi.it

ABSTRACT

This paper analyses artt. 11-13 of law no. 132/2025, highlighting how the Italian legislature views AI as going beyond mere automation, placing it within a framework based on precaution, prevention, transparency and human responsibility. Article 11, which can be characterised as a purpose clause, guides the use of AI within the constitutional framework of artt. 32 and 41, treating algorithmic risk as a factor to be managed even under conditions of uncertainty, in accordance with a precautionary approach. Article 12 introduces a novel form of risk-informed governance, entrusted to an Observatory tasked with contributing to the production of verifiable knowledge as a prerequisite for effective prevention. Finally, art. 13 reaffirms the centrality of human responsibility in the use of AI within the professions, translating transparency into the traceability and accountability of decisions. The analysis shows that, whilst drawing on the framework of the AI Act, the national legislation entrusts the challenge of transforming formal obligations into substantive guarantees to the ability to make risk knowable merely formal obligations.

Keywords: artificial intelligence and work; algorithmic risk; prevention and precaution; transparency and traceability; human responsibility.

<https://doi.org/10.60923/issn.2421-2695/25643>

La legge sulla IA tra finalità, prevenzione e prova nella tutela della salute e sicurezza sul lavoro

SOMMARIO: 1. Premessa – 2. L'art. 11 come norma di scopo: delimitazione funzionale dell'impiego dell'IA nel lavoro. – 2. 1. Trasparenza, informazione e diritto alla spiegazione come presidi di verificabilità dell'impiego dell'IA. – 2.2. Discriminazione algoritmica e rilevanza prevenzionistica del rischio psicosociale. – 2.3. La logica precauzionale nell'art. 11 tra prevenzione e governo dell'incertezza. – 3. L'art. 12 e la *governance* conoscitiva dell'algoritmo: natura, ruolo e limiti dell'Osservatorio. – 4. L'art. 13 e i limiti professionali dell'ausilio algoritmico. – 5. Conclusioni: *dalla dichiarazione alla prova*.

1. Premessa

Nella riflessione scientifica dedicata all'IA, il rischio metodologico più ricorrente consiste nel relegare il fenomeno entro coordinate astratte, popolate da scenari futuribili, promesse di efficienza o timori indeterminati, trascurando invece i contesti concreti nei quali essa incide sull'organizzazione del lavoro, sui processi decisionali e sulle condizioni di tutela della persona.

Se «l'avvento dell'IA generativa rappresenta un salto quantico e non una mera evoluzione tecnologica; essa interviene trasformando i processi decisionali e ponendo problemi radicalmente nuovi che non possono essere trattati come semplici aspetti tecnici»⁽¹⁾, ne discende che «le nuove tecnologie digitali modificano i tempi, gli spazi e i ritmi del lavoro, con riflessi diretti sulla qualità del lavoro e sulla salute e sicurezza»⁽²⁾.

In questa prospettiva, l'intelligenza artificiale entra nel diritto della sicurezza del lavoro non come fenomeno futuribile, ma come tecnologia già normalizzata nei processi organizzativi e decisionali, secondo una traiettoria che trova un primo riconoscimento normativo esplicito anche nella disciplina europea del lavoro mediante piattaforme digitali⁽³⁾, la quale ha anticipato – in ambito settoriale – i temi della gestione algoritmica, della

(1) U. Ruffolo, *Artificial Intelligence Act e AI generativa*, *GI*, 2, 2025, 438 ss.

(2) T. Treu, *Intelligenza Artificiale (IA): integrazione o sostituzione del lavoro umano?*, *CSDLE "Massimo D'Antona".IT*, n. 487/2024. Sulla stessa lunghezza d'onda, tra gli altri, F. Faini, *Intelligenza artificiale e regolazione giuridica: il ruolo del diritto nel rapporto tra uomo e macchine*, *federalismi.it*, 2023, n. 2, 1 ss.; M. Faioli, *Robot Labor Law. Linee di ricerca per una nuova branca del diritto del lavoro e in vista della sessione sull'intelligenza artificiale del G7 del 2024*, *Federalismi.it*, 2024, n. 8, 182 ss.; C. Frascheri, *Gestire il cambiamento*, Edizioni Lavoro, 2023; V. Maio, *Il diritto del lavoro e le nuove sfide della rivoluzione robotica*, *ADL*, 2018, n. 6, 1414 ss.; P. Pascucci, *Le nuove coordinate del sistema prevenzionistico*, *DSL*, 2023, 2, I, 48 ss.; V. Pasquarella, *I principali garanti della sicurezza nell'era delle nuove trasformazioni del lavoro*, *VTDL*, n. 4, 2023, 890 ss.; C. Romeo, *L'era degli algoritmi e la sua incidenza nell'ambito della certezza del diritto: un connubio sospetto*, *LG*, 2024, n. 1, 5 ss.; M. Squeglia, *Obiettivi, strumenti e obiettivi dell'intelligenza artificiale nella gestione della sicurezza del lavoro*, *DSL*, 1, 2025, 114 ss.

(3) Si tratta della Direttiva (UE) 2024/2831 del Parlamento europeo e del Consiglio, del 23 ottobre 2024, relativa al miglioramento delle condizioni di lavoro nel lavoro mediante piattaforme digitali. La

trasparenza sostanziale, della supervisione umana e della responsabilità per le decisioni automatizzate.

L'intervento del legislatore nazionale si colloca lungo questa medesima direzione.

La legge 23 settembre 2025, n. 132 si configura come un intervento di carattere generale volto a governare la ricerca, lo sviluppo e l'impiego dell'IA entro una prospettiva antropocentrica e di tutela dei diritti fondamentali, in coordinamento con il Reg. (UE) 2024/1689 del 13 giugno 2024. In specie, la dimensione lavoristica non assume carattere autonomo, ma si inserisce tra gli ambiti nei quali l'uso dell'IA è idoneo a incidere su beni primari, quali la salute, la sicurezza e l'imputabilità delle decisioni.

Gli artt. 11 e segg., l. n. 132/2025 si collocano esattamente in tale solco: essi individuano, infatti, nella salute, nella sicurezza e nell'organizzazione del lavoro gli ambiti nei quali l'IA manifesta la propria rilevanza giuridica, ossia il punto in cui l'innovazione tecnologica incontra gli obblighi inderogabili di protezione e le coordinate fondamentali della disciplina lavoristica ⁽⁴⁾.

È su questo terreno – quello della prevenzione, della valutazione dei rischi e dei possibili effetti di eterodirezione algoritmica ⁽⁵⁾ – che l'IA rivela, al tempo stesso, la propria forza trasformativa e l'esigenza di un governo giuridico responsabile, in coerenza con le più recenti elaborazioni europee e nazionali.

Da qui prende forma una linea di frattura interpretativa che attraversa una parte significativa della riflessione sull'IA nell'organizzazione del lavoro e, più specificamente, nei suoi riflessi sulla salute e sicurezza dei lavoratori.

direttiva – da recepire entro il 2 dicembre 2026 (art. 29, par. 1) – introduce, in ambito settoriale, obblighi specifici di trasparenza dei sistemi di gestione algoritmica, di supervisione umana effettiva e di responsabilità in relazione alle decisioni automatizzate incidenti sul rapporto di lavoro, anticipando profili e problematiche poi generalizzati dal Regolamento (UE) 2024/1689. In argomento, cfr. V. Bavaro, *Lavoro tramite piattaforme digitali e subordinazione algoritmica*, ADL, 2021, 211-236; A. Rosin - H. Parviainen, *Improving the transparency of algorithmic management in platform work*, *European Labour Law Journal*, 2025, vol. 16(4), 472 ss.; P. Tullini, *La Direttiva piattaforme e i diritti del lavoro digitale*, LLI, 2022, 1, 1 ss.; M. Tomaszewska, *Transparency and Algorithmic Oversight in the Digital Workplace: Legal Challenges under Directive 2024/2831*, in *Studia z Zakresu Prawa Pracy i Polityki Społecznej*, 2025, vol. 32, n. 3, 187 ss.; S. Yannakourou, *Regulating Algorithmic Management of Work at EU Level: The Adequacy of the Platform Work Directive and the AI Act*, in *EU Digital Law in the AI Era*, Springer, 2025, 33-53; L. Zappalà, *Informatizzazione dei processi decisionali e diritto del lavoro: algoritmi, poteri datoriali e responsabilità*, WP CSDLE "Massimo D'Antona", 2021, n. 446.

⁽⁴⁾ Cfr. L. Barbieri, *Sorveglianza umana di un sistema di IA ad alto rischio*, DPL, 46, 2025, 2666, ove si rileva che «l'intervento umano e l'esercizio della sorveglianza costituiscono presupposti essenziali per un utilizzo affidabile dell'IA, che deve restare uno strumento nel rispetto della dignità umana e dell'autonomia personale».

⁽⁵⁾ È convincimento comune in dottrina che l'automazione dei processi decisionali introduce, infatti, una forma inedita e più sottile di eterodirezione: non più esercitata direttamente dal datore di lavoro, ma mediata da correlazioni, inferenze statistiche e procedure algoritmiche che incidono sull'organizzazione e sulle modalità concrete di svolgimento della prestazione. Tale dinamica amplia l'area del potere datoriale anche in assenza di una intenzionalità soggettiva, poiché l'azione dell'algoritmo è idonea a orientare comportamenti, determinare assegnazioni, modulare ritmi e valutare *performance* secondo logiche autonome rispetto alla razionalità umana e, per ciò stesso, bisognose di controllo, trasparenza e verifica.

Da un lato, vi è chi ritiene che la penetrazione dell'IA nei processi organizzativi e decisionali non consenta di relegare la tecnologia ad un mero ausilio tecnico, poiché essa incide direttamente sulle coordinate del paradigma lavoristico e sollecita una revisione delle categorie fondamentali – dall'imputazione della volontà al nesso di responsabilità, dalla nozione di controllo alla vigilanza datoriale ⁽⁶⁾.

Dall'altro lato, pur muovendo dalla forza cogente del rinvio al Reg. (UE) 2024/1689 (*AI Act*) – che tutela i diritti fondamentali e introduce obblighi di tracciabilità, documentazione e supervisione umana – si è affermata una lettura che tende a ricondurre tali prescrizioni entro una dimensione prevalentemente tecnica e regolatoria, coerente con l'approccio *risk-based* del legislatore europeo, attenuandone l'incidenza sistemica sulle categorie proprie del diritto del lavoro e della sicurezza, e rinviando agli ordinamenti settoriali il compito di tradurre le garanzie formali in tutele effettive ⁽⁷⁾.

In questo spazio di tensione emerge una ulteriore prospettiva ⁽⁸⁾ che, senza negare la rilevanza di entrambe le letture, propone di assumere come punto di osservazione privilegiato non la mera incidenza dell'IA sul lavoro in quanto tale, bensì gli effetti che il suo impiego produce sulla salute e sicurezza dei lavoratori, quali dimensioni costituzionalmente protette ai sensi dell'art. 32 Cost. e solo in via mediata riconducibili all'esercizio della libertà di iniziativa economica *ex art.* 41 Cost.

Una ricostruzione che l'intervento del legislatore nazionale consente ora di valorizzare, posto che gli artt. 11 e segg., l. n. 132/2025 – letti congiuntamente all'obbligo

⁽⁶⁾ Non è possibile dar conto della vasta bibliografia che converge su tale impostazione; senza alcuna pretesa di esaustività, e con rinvio ai contributi ivi citati, cfr. S. Cairoli, *Intelligenza artificiale e sicurezza sul lavoro: uno sguardo oltre la siepe*, *DSL*, 2, 2024; S. Ciucciovino, *Intelligenza artificiale e diritto del lavoro: problemi e prospettive - Risorse umane e intelligenza artificiale alla luce del regolamento (UE) 2024/1689, tra norme legali, etica e codici di condotta*, *DRI*, 3, 2024; M. G. Elmo, *Sistemi IA e rischi per la salute e la sicurezza dei lavoratori: riflessioni a margine della regolamentazione europea*, *AmbienteDiritto.it*, XXIV, 4, 2024; S. Marassi, *Intelligenza artificiale e sicurezza sul lavoro*, in M. Biasi (a cura di), *Diritto del lavoro e intelligenza artificiale*, Giuffrè, 2024, 207 ss.; P. Pascucci - C. Lazzari, *Sistemi di IA, salute e sicurezza sul lavoro: una sfida al modello di prevenzione aziendale, fra responsabilità e opportunità DSL*, 4, 2024, 587 ss.; M. Peruzzi, *Sistemi automatizzati e tutela della salute e sicurezza sul lavoro*, *DSL*, 2024, n. 2, 78 ss. Una convergenza analoga si rinviene, anche al di fuori del paradigma strettamente lavoristico e della sicurezza del lavoro, in contributi che colgono il carattere qualitativamente nuovo dell'autonomia algoritmica e le sue ricadute sui tradizionali schemi di imputazione, controllo e responsabilità: N. Cristianini, *Machina sapiens. L'algoritmo che ci ha rubato il segreto della conoscenza*, Il Mulino, 2024, spec. 112 ss.; L. Floridi, *AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models*, in *Philosophy & Technology*, 2023, 36(15); U. Ruffolo, *Artificial Intelligence Act e AI generativa*, cit., 439 ss.; nonché, dello stesso Autore, *Per i fondamenti di un diritto della robotica self-learning: dalla machinery produttiva all'auto driverless: verso una «responsabilità da algoritmo»?», in Intelligenza artificiale e responsabilità*, Giuffrè, 2017, 1 ss.

⁽⁷⁾ In ordine, invece, alla lettura che riconduce l'IA prevalentemente entro una logica di regolazione tecnico-prevenzionistica e di gestione del rischio, fondata sull'approccio *risk-based* adottato dall'*AI Act*, si v. P. Loi, *Il rischio proporzionato nella proposta di regolamento sull'IA*, in *federalismi.it*, 2023, 4, 241 ss.; con particolare riferimento ai profili di implementazione del Regolamento e alla letteratura di matrice *risk-regulatory*, v. inoltre C. Novelli, *L'Artificial Intelligence Act europeo: alcune questioni di implementazione*, *federalismi.it*, 2024, 2, 96 ss.

⁽⁸⁾ V., già, M. Squeglia, *Obiettivi, strumenti e metodi dell'intelligenza artificiale*, cit., spec. 126 ss.

generale di protezione di cui all'art. 2087 c.c. – non si limitano a recepire l'impostazione eurounitaria, ma risultano idonei a operare, sul piano sistemico, quali parametri potenzialmente rilevanti ai fini dell'imputazione della responsabilità, nella misura in cui delimitano le condizioni di legittimità dell'esercizio del potere decisionale assistito o condizionato da sistemi algoritmici.

La premessa normativa così delineata consente di individuare il baricentro della riforma: l'intelligenza artificiale non è trattata come dispositivo di mera automazione, ma come oggetto giuridico che deve lasciare tracce documentabili e rimanere riconducibile alla responsabilità umana. Ne discende il passaggio dalla trasparenza meramente proclamata ad una "trasparenza dimostrabile", come evidenziato dal D.M. 17 dicembre 2025, n. 180, in cui la documentazione assolve a una funzione probatoria essenziale ⁽⁹⁾.

L'interrogativo preliminare attiene, dunque, alle condizioni e alle modalità della decisione di procedere all'adozione di un sistema di intelligenza artificiale poiché – come previsto dal Reg. (UE) 2024/1689 – questi ultimi impiegati nell'ambito dell'occupazione e della gestione dei lavoratori, laddove incidano su decisioni rilevanti per il rapporto di lavoro, rientrano nel perimetro dei sistemi qualificabili come "ad alto rischio", con l'effetto di attivare, in capo al soggetto utilizzatore (il cd. *deployer*), gli obblighi speciali disciplinati dal Capo III ⁽¹⁰⁾.

Proprio in ragione di tale qualificazione normativa, la scelta di introdurre un sistema di IA non può essere ricondotta ad una mera opzione tecnologica o organizzativa, rimessa a valutazioni intuitive o a logiche di convenienza contingente. Essa richiede, al contrario, una fase istruttoria rigorosa, nella quale il soggetto utilizzatore sia chiamato a esaminare in modo critico la documentazione tecnica disponibile, le istruzioni d'uso, le caratteristiche funzionali e i limiti operativi dei sistemi esistenti, al fine di individuare la soluzione coerente con le finalità perseguite e con il concreto contesto organizzativo.

Tale esigenza trova fondamento nei più generali obblighi di organizzazione e controllo posti dall'art. 2086, co. 2, c.c., che impongono all'impresa di adottare assetti adeguati alla natura e alle dimensioni dell'attività, idonei a sorreggere decisioni fondate su processi valutativi strutturati, razionalmente giustificabili e documentabili. L'eventuale adozione di un sistema di IA rientra, dunque, tra le decisioni organizzative qualificate che presuppongono un percorso valutativo tracciabile e motivato, volto a dimostrare la

⁽⁹⁾ Le Linee guida ministeriali precisano che «la trasparenza implica una tracciabilità effettiva delle logiche algoritmiche e richiede un controllo umano sostanziale», qualificandola espressamente come condizione di verificabilità (cfr. L. Barbieri, *Sistemi di LA nel mondo del lavoro: Linee guida ministeriali ed elementi di raccordo con la normativa vigente*, DPL, 28, 2025, 1710 ss.; sul tema di trasparenza, si v. anche F. Costantini, *La L. n. 132/2025 sull'intelligenza artificiale: una prima lettura e alcune indicazioni in materia di diritto del lavoro*, LG, 12, 2025, 1093 ss.)

⁽¹⁰⁾ S. Ciucciovino, *Intelligenza artificiale e diritto del lavoro: problemi e prospettive - Risorse umane e intelligenza artificiale alla luce del regolamento (UE) 2024/1689, tra norme legali, etica e codici di condotta*, DRI, 3, 2024.

congruità tecnica e funzionale della tecnologia prescelta rispetto agli obiettivi aziendali e alle garanzie ordinamentali coinvolte ⁽¹¹⁾.

Solo all'esito di tale fase preliminare può essere affrontata la disciplina di cui alla l. n. 132/2025, poiché è in questo passaggio che la questione si ridefinisce nei suoi termini essenziali: non se introdurre sistemi di intelligenza artificiale, ma come governarne l'uso all'interno dell'organizzazione ⁽¹²⁾.

Sotto questo profilo, l'art. 11 rappresenta il primo tassello della disciplina, poiché ne definisce finalità e architettura funzionale e consente di cogliere il tentativo di rendere praticabile la logica unitaria che orienta l'intero intervento legislativo: un modello nel quale prevenzione, prova, responsabilità umana e governo del rischio algoritmico convergono in un unico asse di protezione giuridica.

2. L'art. 11 come norma di scopo: delimitazione funzionale dell'impiego dell'IA nel lavoro

Il co. 1 dell'art. 11, l. n. 132/2025 si configura come una norma di scopo: non definisce una fattispecie tipica, né individua un modello operativo chiuso, ma enuncia le finalità e i principi cui deve conformarsi l'impiego dell'intelligenza artificiale nei rapporti di lavoro. In altri termini, la disposizione delimita il perimetro finalistico dell'utilizzo dell'intelligenza artificiale, precisandone le condizioni teleologiche di ammissibilità. Essa, tuttavia, non fonda in sé l'impiego della tecnologia, ma ne circoscrive la sfera funzionale,

⁽¹¹⁾ In dottrina è stato osservato che l'adeguatezza degli assetti organizzativi *ex art.* 2086, co. 2, c.c. non si esaurisce in adempimenti meramente formali, ma implica la capacità dell'impresa di «tradurre il rischio in decisione organizzativa strutturata», rendendo tracciabili, verificabili e motivabili le scelte assunte in funzione preventiva (M. Faioli, *Adeguatezza ex art. 2086 c.c. e obbligo di introdurre tecnologia avanzata per mitigare i rischi da lavoro*, *federalismi.it*, 2025, 6, 143 ss.). Muovendo da tale impostazione, l'A. afferma che, in determinate circostanze, l'adeguatezza degli assetti possa «coincidere con l'obbligo di introdurre tecnologia avanzata, anche nella forma di intelligenza artificiale, per mitigare i rischi da lavoro». Tale affermazione, pur cogliendo correttamente il nesso tra governo del rischio e assetti organizzativi, richiede tuttavia una lettura non automatica e non deterministica. L'art. 2086, co. 2, c.c. non sembra infatti imporre un obbligo generalizzato di innovazione tecnologica, né consente di assumere la mancata introduzione dell'intelligenza artificiale come indice immediato di inadeguatezza dell'assetto. L'oggetto dell'obbligo non è la tecnologia in sé, ma il processo decisionale che conduce alla sua eventuale adozione. In questa prospettiva, la tecnologia avanzata non costituisce parametro autonomo di adeguatezza, ma oggetto della valutazione organizzativa richiesta dalla norma: ciò che rileva è la capacità dell'impresa di dimostrare che l'adozione (o la mancata adozione) di un sistema di intelligenza artificiale sia il risultato di una valutazione preventiva, proporzionata e documentata, svolta in relazione ai rischi concretamente presenti, agli strumenti disponibili e alla struttura dell'organizzazione. Una lettura dell'art. 2086 c.c. che individui nell'IA un obbligo *in re ipsa* finirebbe, altrimenti, per spostare l'asse dell'adeguatezza dal governo responsabile del rischio alla mera sofisticazione dello strumento, con il rischio di attenuare, anziché rafforzare, la responsabilità organizzativa dell'impresa.

⁽¹²⁾ Sull'idea della sicurezza del lavoro come dimensione eminentemente organizzativa, e non riducibile ad un insieme di presidi o adempimenti meramente tecnici, specie nei contesti di progressiva dematerializzazione degli spazi lavorativi: cfr. C. Lazzari, *Lavoro senza luogo fisso, dematerializzazione degli spazi, salute e sicurezza*, *LLI*, 2023, n. 1, 22 ss.

rinviano alle fonti che disciplinano in modo effettivo obblighi, limiti e garanzie relativi al suo utilizzo.

È a tali fonti che occorre guardare per identificare i presupposti di liceità dell'IA nel lavoro: agli obblighi prevenzionistici sanciti dal d.lgs. n. 81/2008; alla disciplina in materia di protezione dei dati personali dettata dal Reg. (UE) 2016/679 e dal d.lgs. n. 196/2003, come modificato dal d.lgs. n. 101/2018; nonché al quadro tecnico e regolatorio specifico delineato dal Reg. (UE) 2024/1689, che definisce requisiti, presidi e obblighi per i sistemi di IA, in particolare quando siano impiegati nei contesti occupazionali ⁽¹³⁾.

Ebbene, si tratta di una disposizione la cui tecnica normativa richiama quella delle norme preliminari del Codice della Navigazione ⁽¹⁴⁾. Come l'art. 1 di quel codice, anche l'art. 11, l. n. 132/2025 non è destinato a produrre effetti autonomi, ma richiede, per assumere operatività, l'intervento di altre fonti legislative che ne attuino il contenuto. La norma, pertanto, non disciplina direttamente l'impiego dell'IA nel rapporto di lavoro: ne definisce la finalità e ne circoscrive l'area funzionale, lasciando agli altri plessi normativi il compito di conformare obblighi, limiti e condizioni di liceità.

Se si condivide che il primo co. dell'art. 11 è norma di scopo, il discrimine applicativo relativo ai possibili impieghi dell'intelligenza artificiale deve essere filtrato tramite il rinvio alla fonte costituzionale.

Quando la tecnologia è introdotta per proteggere o migliorare l'integrità psicofisica del lavoratore, il parametro di riferimento è l'art. 32 Cost. in connessione con l'art. 2087 c.c.: l'intelligenza artificiale opera come misura di prevenzione rispetto a rischi fisici, organizzativi e psicosociali, senza interferire con la prestazione individuale. In questa prospettiva, ciò che rileva è l'effetto: migliorare il contesto prevenzionistico, non valutare o sorvegliare il singolo ⁽¹⁵⁾.

Se, invece, la tecnologia è introdotta per accrescere efficienza, qualità o produttività, il parametro diviene l'art. 41 Cost.: la finalità è legittima, ma incontra il limite della sicurezza, della libertà e della dignità della persona. Quando lo strumento osserva, inferisce o valuta comportamenti e prestazioni individuali, esso diventa idoneo al controllo e si attiva l'art. 4, l. n. 300/1970, con i presidi di legalità, necessità, proporzionalità e trasparenza ⁽¹⁶⁾.

⁽¹³⁾ In senso parzialmente critico rispetto a tale estensione, v. F. Costantini, *La legge n. 132/2025 e l'intelligenza artificiale nei rapporti di lavoro*, cit., 1093.

⁽¹⁴⁾ A mente del quale stabilisce: "In materia di navigazione, marittima, interna ed aerea, si applicano il presente codice, le leggi, i regolamenti, le norme corporative e gli usi ad essa relativi. Ove manchino disposizioni del diritto della navigazione e non ve ne siano di applicabili per analogia, si applica il diritto civile".

⁽¹⁵⁾ È stato osservato che «i processi decisionali modificano criteri di valutazione e forme di controllo, potendo essere condizionati da discriminazioni algoritmiche determinate da un anomalo funzionamento del sistema di IA» (L. Barbieri, *Sistemi di LA nel mondo del lavoro*, cit., 1712).

⁽¹⁶⁾ In questo scenario rientrano le applicazioni che interagiscono direttamente con la prestazione lavorativa: sistemi che assegnano compiti sulla base dell'efficienza pregressa, algoritmi che valutano i tempi di risposta o individuano errori ricorrenti, nonché modelli predittivi della produttività individuale. In tali ipotesi emerge una capacità – anche non intenzionale – di monitoraggio, idonea ad attivare l'ambito applicativo dell'art. 4 St. lav., poiché incide sull'invasività, sull'intensità e sulla utilizzabilità della sorveglianza

In altri termini: non rileva lo scopo dichiarato, ma l'effetto concreto sull'organizzazione e sulla persona ⁽¹⁷⁾.

Una tale ricostruzione non resta sul piano teorico, ma produce effetti funzionali immediati.

Nel binario dell'art. 32 Cost., in combinazione con l'art. 2087 c.c., la valutazione dei rischi di cui all'art. 28, d.lgs. n. 81/2008 deve ricomprendere anche i rischi psicosociali derivanti dall'adozione di tecnologie, e impone di governarli mediante un sistema informativo e formativo adeguato, nonché attraverso il rispetto delle regole sul trattamento dei dati personali, quali informazione strutturata, minimizzazione, eventuale valutazione d'impatto, intervento umano nei processi decisionali automatizzati.

Nel binario dell'art. 41, quando l'impiego dell'IA risulta idoneo – in ragione delle sue modalità tecniche – a osservare, inferire o valutare comportamenti e prestazioni individuali, si attiva l'art. 4, l. n. 300/1970: la valutazione dei rischi “algoritmici” deve allora includere ritmi, carichi cognitivi, *stress* e profili valutativi della *performance*, mentre il rispetto congiunto del Reg. (UE) 2016/679 e del Reg. (UE) 2024/1689 impone documentazione coerente, registri accurati, tracciabilità ricostruttiva ed una supervisione umana effettiva, in grado di confermare, correggere o arrestare il sistema.

In questo bilanciamento tra gli artt. 32 e 41 si misura la liceità sostanziale dell'impiego dell'IA: non nella dichiarazione programmatica della finalità perseguita, ma nella prova della sua conformità ai parametri costituzionali che presidiano la persona nel lavoro.

2.1. Trasparenza, informazione e diritto alla spiegazione come presidi di verificabilità dell'impiego dell'IA

Nel quadro di un'articolazione teleologicamente orientata, i commi successivi dell'art. 11 concorrono a definire le condizioni di attuazione dello scopo enunciato dal co. 1, delineando un sistema che salda obblighi informativi, garanzie di verificabilità e limiti all'esercizio del potere datoriale.

nei confronti del lavoratore, secondo modalità profondamente diverse da quelle per cui la norma era stata originariamente concepita, vale a dire il controllo “a distanza”. In questa prospettiva, il potenziale di controllo risulta incorporato nello strumento organizzativo stesso, rendendo strutturale la possibilità di osservazione della prestazione. Su questo specifico profilo, v. F. Rotondi - A. Tursi, *I controlli sull'attività dei lavoratori alla prova dell'intelligenza artificiale*, *LG*, 2, 2025, 113 ss. In dottrina, cfr. I. Alvino, *I nuovi limiti al controllo a distanza dell'attività dei lavoratori nell'intersezione fra le regole dello Statuto dei lavoratori e quelle del Codice della privacy*, *LLI*, vol. 2, 1, 2016, 35 ss.; E. Gragnoli, *I controlli sul lavoratore e i cosiddetti sistemi di intelligenza artificiale*, *LDE*, 2024, n. 3, 1 ss.; V. Maio, *Il diritto del lavoro e le nuove sfide della rivoluzione robotica*, *ADL*, 2018, n. 6, 1414 ss.; P. Tullini, *Controlli a distanza e tutela dei dati personali del lavoratore*, Giappichelli, 2017; C. Zoli, *Il controllo a distanza dell'attività dei lavoratori e la nuova struttura dell'articolo 4 della legge 300/1970*, *VTDL*, 4, 2016.

⁽¹⁷⁾ Sul rischio che i modelli generativi producano “*output imprevedibili*”, con effetti discriminatori o distorsivi, v. U. Ruffolo, *Artificial Intelligence Act e AI generativa*, cit., 439 ss., ove si richiamano le “proprietà emergenti” dei LLMs (*Large Language Models*), e la loro capacità di svolgere compiti non previsti in addestramento.

Il co. 2 dispone che «(...) L'utilizzo dell'intelligenza artificiale in ambito lavorativo deve essere sicuro, affidabile, trasparente e non può svolgersi in contrasto con la dignità umana né violare la riservatezza dei dati personali. Il datore di lavoro o il committente è tenuto a informare il lavoratore dell'utilizzo dell'intelligenza artificiale nei casi e con le modalità di cui all'articolo 1-*bis* del decreto legislativo 26 maggio 1997, n. 152».

In tale assetto, la trasparenza assume la natura di obbligo probatorio: non una mera comunicazione informativa, ma una condizione di verificabilità sostanziale del sistema e, dunque, della legittimità del processo decisionale. Una lettura che circoscriva l'obbligo informativo ai soli lavoratori non coglie – a nostro avviso – la portata sistematica del rinvio operato dall'art. 11.

L'art. 1-*bis*, d.lgs. n. 152/1997 si innesta, infatti, entro una costruzione costituzionale a due livelli cui la l. n. 132/2025 si limita semplicemente a rinviare.

Da un lato, il richiamo all'art. 1-*bis* comporta – per il tramite del primo comma – che il datore di lavoro debba rendere un'informazione strutturata e completa sulla logica, sulle finalità e sulle modalità di funzionamento dei sistemi decisionali o di monitoraggio integralmente automatizzati, così da permettere alla persona che lavora di comprendere in modo effettivo l'impatto che tali sistemi possono produrre sulla prestazione e sull'organizzazione del lavoro.

Dall'altro lato, lo stesso rinvio determina – per il tramite del terzo comma del medesimo art. 1-*bis* – un ampliamento della legittimazione soggettiva all'accesso: alle rappresentanze sindacali aziendali o territoriali è attribuito un diritto autonomo di richiesta delle medesime informazioni, esercitabile indipendentemente dall'iniziativa individuale e sempre entro il perimetro dei sistemi richiamati dal co. 1 ⁽¹⁸⁾.

Ne risulta un modello informativo unitario, privo di duplicazioni o distinzioni contenutistiche, che affianca alla tutela individuale un presidio collettivo capace di verificare l'effettivo impiego dei sistemi di IA, la coerenza tra finalità dichiarate e configurazioni operative e gli effetti complessivi che tali tecnologie producono sull'assetto organizzativo.

Sul versante dell'art. 32, in combinazione con l'art. 2087 c.c., l'informazione diventa elemento strutturale della prevenzione: incide sulla valutazione dei rischi, sulla consultazione, sulla formazione e sulle misure tecniche e organizzative, permettendo di ricondurre entro un quadro conoscitivo minimo l'opacità e l'adattività proprie dei sistemi algoritmici.

⁽¹⁸⁾ Non è possibile ricavare dall'art. 1-*bis*, co. 3, d.lgs. n. 152/1997 una differenziazione tale da introdurre un regime informativo distinto rispetto a quello delineato dai commi 1 e 2. La disposizione non prevede, infatti, alcuna differenza qualitativa tra le informazioni che il datore di lavoro è tenuto a fornire al lavoratore e quelle che possono essere richieste dalle rappresentanze sindacali: il comma 3 attribuisce semplicemente alle RSU/RSA un canale autonomo – e non subordinato alla richiesta individuale – per accedere al medesimo complesso informativo definito dal comma 2 (*lett. a-f*). Si tratta, pertanto, di una distinzione soggettiva e non oggettiva, finalizzata a garantire una verifica collettiva sull'impiego dei sistemi decisionali o di monitoraggio integralmente automatizzati, senza che ciò modifichi né l'oggetto né la portata dell'obbligo informativo.

Per contro, sul versante dell'art. 41, essa integra le condizioni di legittimità per l'impiego di tecnologie che – per le loro caratteristiche tecniche – risultano normalmente idonee a incidere sull'esercizio del potere di organizzazione e, in via strutturale, anche su quello di controllo. Ciò esige un flusso informativo scritto, completo e ricostruibile e, ove lo strumento presenti capacità di osservazione o inferenza rispetto alla prestazione, l'attivazione della procedura sindacale o autorizzatoria prescritta dall'art. 4 St. lav.

In questa prospettiva, la fase istruttoria preliminare e l'obbligo informativo successivo non rappresentano momenti separati, ma segmenti di una medesima architettura funzionale, volta a rendere ricostruibile il rischio e valutabili *ex ante* i profili tecnici, organizzativi e cognitivi dei sistemi adottati.

Ciò vale anche nel caso dei sistemi di intelligenza artificiale qualificati come “ad alto rischio” dal Reg. (UE) 2024/1689, poiché il differimento dell'applicazione del Capo III non incide sugli obblighi prevenzionistici già vigenti nell'ordinamento interno, né elimina l'esigenza di un controllo organizzativo, documentale e probatorio effettivo e tracciabile ⁽¹⁹⁾.

Proprio in questa direzione, il co. 2 dell'art. 11 non introduce obblighi autonomi, ma si coordina con il sistema probatorio delineato dal Reg. (UE) 2024/1689 – fondato su documentazione tecnica, tracciabilità dei processi, registrazione delle operazioni e supervisione umana effettiva – e con le Linee guida adottate con il D.M. 17 dicembre 2025, n. 180, che includono espressamente i rischi psicosociali e organizzativi tra le principali aree di attenzione nella *governance* dei sistemi di IA ⁽²⁰⁾.

Ebbene, da quanto sin qui esposto, il rinvio operato dal legislatore all'art. 1-*bis* ripropone perplessità già manifestatesi, tanto in sede di primo commento alla disciplina introdotta dal c.d. decreto trasparenza del 2022, quanto con riferimento al d.d.l. della XIX Legislatura, presentato al Senato il 20 maggio 2024 (A.S. n. 1146) e poi confluito nella l. n. 132/2025, laddove era stato segnalato il rischio che l'opzione per un rinvio alla disciplina dell'automazione tradizionale non fosse idonea a cogliere la specificità tecnologica e giuridica dell'IA ⁽²¹⁾. Tali criticità risultano oggi ancor più evidenti alla luce della diffusione

⁽¹⁹⁾ Com'è noto, nel Reg. (UE) 2024/1689 i sistemi di IA sono qualificati come “ad alto rischio” quando rientrano nelle categorie individuate dal Capo III e dall'Allegato III, in particolare ove incidano su accesso al lavoro, gestione del rapporto, valutazioni, selezioni o licenziamenti. Gli obblighi sostanziali previsti per tali sistemi – tra cui la gestione del rischio, la documentazione tecnica, la tenuta dei registri e la supervisione umana – sono soggetti ad un regime di applicazione progressiva, operando in via generale a decorrere dall'entrata in applicazione del Capo III. Resta però fermo che la qualificazione come “sistema ad alto rischio” rileva in ogni caso quale parametro di orientamento nella valutazione degli obblighi prevenzionistici, organizzativi e informativi derivanti dall'ordinamento interno e dalla normativa nazionale di coordinamento, anche anteriormente alla piena operatività del quadro europeo.

⁽²⁰⁾ Alcuni A. hanno evidenziato che «non sempre l'informazione prescritta è in grado di dischiudere la logica dei sistemi impiegati, specie quando operano secondo schemi non immediatamente riconducibili alla razionalità umana», generando *opacità strutturale* (L. Barbieri, *Sistemi di LA nel mondo del lavoro*, cit., 1713 ss.).

⁽²¹⁾ M. Squeglia, *Obiettivi, strumenti e metodi dell'intelligenza artificiale*, cit., 129, nota n. 50.

di forme di governo algoritmico del lavoro caratterizzate da un'elevata complessità tecnica e decisionale.

Il nodo problematico si manifesta soprattutto nelle ipotesi in cui l'automazione incida in modo sostanziale sul processo decisionale pur senza tradursi, sul piano formale, in una decisione integralmente automatizzata. In tali casi, l'*output* del sistema può operare in assenza di un controllo umano effettivo e, nondimeno, restare al di fuori della definizione normativa di "sistema di IA" delineata dal Reg. (UE) 2024/1689, determinando uno scarto regolativo tra la struttura delle garanzie e la realtà funzionale delle tecnologie impiegate ⁽²²⁾.

È in questo spazio intermedio – quello delle decisioni parzialmente automatizzate o ibride – che si concentra l'area di maggiore criticità del governo algoritmico del lavoro: sistemi che, pur mantenendo un intervento umano sul piano dichiarato, possono risultare automatizzati sul piano sostanziale, laddove l'apporto della persona si risolva in una mera ratifica dell'*output* del sistema, con conseguente attenuazione dell'effettività degli obblighi informativi e delle possibilità di presidio dei rischi per i diritti del lavoratore. Cosicché per superare questa rigidità dicotomica che alcuni Autori hanno proposto una lettura graduata dell'obbligo informativo, idonea a correlare l'intensità degli adempimenti non già alla formale qualificazione del sistema come integralmente automatizzato, bensì al livello di complessità tecnologica e di incidenza effettiva sul processo decisionale ⁽²³⁾, secondo una progressione che consente di ricondurre l'art. 1-*bis* entro una logica coerente di governo del rischio algoritmico.

Ma, anche a voler ritenere superato il profilo della mera conoscibilità tecnica del sistema, l'informazione non può esaurirsi in un dato descrittivo o meramente ricognitivo: essa deve piuttosto essere accompagnata da un'effettiva spiegazione del funzionamento del sistema.

È proprio su questo terreno che, a nostro avviso, prende forma un nodo problematico meritevole di particolare attenzione. In mancanza di prescrizioni normative puntuali, l'ordinamento sembra tuttavia offrire una possibile via compensativa: il ricorso all'autonomia contrattuale collettiva, in quanto funzionale alla costruzione di un vero e proprio diritto alla spiegazione dei processi decisionali automatizzati, cui si accompagna il correlato obbligo datoriale di assicurarne un'effettiva fruibilità.

⁽²²⁾ In questa direzione ha inciso, in senso restrittivo, anche l'interpretazione fornita dal Ministero del lavoro con circ. n. 19 del 20 settembre 2022, la quale, nel chiarire l'ambito applicativo dell'art. 1-*bis* del d.lgs. n. 152/1997, ha circoscritto gli obblighi informativi ai soli sistemi "integralmente automatizzati", valorizzando in modo rigidamente letterale il dato normativo.

⁽²³⁾ In particolare, si è osservato che, a seguito del rinvio dell'art. 11, l. n. 132/2025 all'art. 1-*bis*, d.lgs. n. 152/1997, possono individuarsi almeno quattro livelli di complessità crescente delle tecnologie impiegate nel rapporto di lavoro: dai trattamenti digitali ordinari, ai sistemi automatizzati con incidenza organizzativa, ai sistemi decisionali o di monitoraggio automatizzati, fino ai sistemi di intelligenza artificiale (specie ad alto rischio), cui corrispondono adempimenti informativi progressivamente più intensi in capo al datore di lavoro. Cfr. F. Costantini, *La legge n. 132/2025 sull'intelligenza artificiale: una prima lettura*, cit., 1095 ss.

D'altra parte, la contrattazione collettiva si configura come la sede elettiva per colmare le asimmetrie informative e per introdurre *standard* di trasparenza più elevati rispetto al minimo legislativo, consentendo così una regolazione effettiva e contestualizzata delle garanzie connesse all'impiego dell'IA nei rapporti di lavoro ⁽²⁴⁾.

Le prime analisi sui rinnovi dei CCNL del 2023-2025 mostrano che la contrattazione si è già orientata verso clausole dedicate all'uso dell'IA, includendo obblighi di informazione rafforzati, valutazioni d'impatto, supervisione umana e modelli di *governance* algoritmica condivisa ⁽²⁵⁾.

È stato osservato in dottrina che la trasparenza algoritmica non può essere garantita soltanto attraverso gli obblighi di fonte legale e che, in assenza di strumenti idonei a disvelare la logica dei sistemi, è proprio la contrattazione collettiva a dover assumere un ruolo centrale, introducendo forme di "trasparenza partecipativa" capaci di riequilibrare il rapporto tra datore di lavoro e lavoratore nella gestione algoritmica dei processi decisionali.

Ne deriva che la previsione contrattuale di un diritto alla spiegazione non costituisce una deviazione rispetto all'assetto vigente, bensì un esercizio fisiologico della capacità regolativa delle parti sociali: una capacità che si rivela oggi tanto più essenziale quanto più l'opacità algoritmica, la diversa razionalità dei sistemi e l'ampliamento del potere datoriale rendono necessario un controllo collettivo non meramente formale, ma sostanziale.

A ben guardare, l'ordinamento sembra offrire una seconda via di compensazione, rappresentata dalla partecipazione dei lavoratori alla gestione dell'impresa, come recentemente disciplinata dalla l. 15 maggio 2025, n. 76. In particolare, la partecipazione gestionale e, soprattutto, quella organizzativa e consultiva si configurano come sedi pienamente idonee a cristallizzare, anche su base volontaria, un diritto alla spiegazione algoritmica. La legge consente, nelle imprese che lo prevedano sul piano statutario, il coinvolgimento dei lavoratori negli organi societari (partecipazione gestionale), nonché l'istituzione di commissioni paritetiche o di organismi consultivi chiamati a intervenire su profili organizzativi e produttivi (partecipazione organizzativa e consultiva).

Tali spazi possono, inoltre, instaurare un dialogo diretto con gli attori istituzionali della prevenzione, quali il rappresentante dei lavoratori per la sicurezza, il medico competente e il preposto: figure che, per le funzioni loro proprie, risultano già collocate nelle aree deputate alla valutazione e al governo dei rischi ⁽²⁶⁾.

⁽²⁴⁾ E. Ghera, *L'articolo 39 della Costituzione e il contratto collettivo*, WP C.S.D.L.E. "Massimo D'Antona", n. 202/2014.

⁽²⁵⁾ P. De Petris, *La questione della trasparenza algoritmica nei rapporti di lavoro interessati dal management automatizzato: una lettura integrata*, MGL, 4, 2024, 781 ss.; A. Ingrao, *Intelligenza artificiale nei rinnovi dei contratti collettivi nazionali di categoria. Una prima roadmap*, LDE, 27 ottobre 2025.

⁽²⁶⁾ In specie il Rappresentante dei Lavoratori per la Sicurezza, titolare di poteri consultivi e di accesso alle informazioni rilevanti per la sicurezza, può essere integrato – nelle commissioni organizzative o consultive previste dalla legge n. 76/2025 – quale destinatario privilegiato delle informazioni sulla logica, i parametri e gli impatti degli algoritmi impiegati nei processi lavorativi. Il medico competente, quale garante della sorveglianza sanitaria e della valutazione dei rischi psicosociali, è in posizione idonea a valutare gli effetti che sistemi non spiegabili, opachi o discriminatori possono generare sul benessere psichico e sulle condizioni di salute dei lavoratori. Il preposto, infine, responsabile dell'attuazione delle misure di sicurezza

In questa prospettiva, la via partecipativa prevista dalla l. n. 76/2025 si salda direttamente con il modello prevenzionistico del d.lgs. n. 81/2008, consentendo di cristallizzare un diritto alla spiegazione algoritmica non solo come presidio di trasparenza, ma come componente dell'analisi e gestione dei rischi psicosociali. Le commissioni partecipative possono cioè costituire la sede di raccordo operativo tra gli obblighi informativi sull'uso dell'IA, il ruolo degli attori della sicurezza e la valutazione dei rischi *ex art. 28*, dando concretezza all'esigenza di supervisione umana effettiva e di controllo sostanziale sui processi decisionali automatizzati. Tali sedi possono includere – senza alterare il sistema normativo – obblighi di informativa algoritmica, procedure di *audit*, diritti di accesso alle logiche di funzionamento dei sistemi e meccanismi di supervisione umana.

2.2. Discriminazione algoritmica e rilevanza prevenzionistica del rischio psicosociale

Nel prosieguo dell'analisi dell'art. 11, l. n. 132/2025, merita attenzione il co. 3, il quale completa il quadro normativo stabilendo che «l'intelligenza artificiale, nell'organizzazione e nella gestione del rapporto di lavoro, deve garantire il rispetto dei diritti inviolabili del lavoratore e l'assenza di discriminazioni, dirette o indirette, fondate sul sesso, sull'età, sulle origini etniche, sul credo religioso, sull'orientamento sessuale, sulle opinioni politiche e sulle condizioni personali, sociali ed economiche, in coerenza con il diritto dell'Unione».

La clausola positivizza così il vincolo antidiscriminatorio come condizione intrinseca di liceità dell'impiego dell'intelligenza artificiale nei rapporti di lavoro, collocandolo non già sul piano meramente correttivo o sanzionatorio, bensì quale limite strutturale all'esercizio del potere organizzativo e gestionale datoriale.

Ma la portata del co. 3 non si esaurisce nel divieto di discriminazione.

Un algoritmo che discrimina – o che genera effetti differenziati senza intenzionalità – produce inevitabilmente un impatto sulla salute e sulla sicurezza, integrando un “rischio psicosociale” rilevante alla stregua dell'art. 28, d.lgs. n. 81/2008 ⁽²⁷⁾.

Si è affermato che i sistemi automatizzati, specie se fondati su modelli statistico-probabilistici o di apprendimento automatico, possono incorporare *bias* che alterano valutazioni, mansioni, carichi di lavoro e percorsi professionali, con effetti non prevedibili *ex ante* ⁽²⁸⁾.

e dell'organizzazione del lavoro, è un interlocutore naturale nella gestione degli strumenti algoritmici che incidono sulla distribuzione delle mansioni, dei carichi e dei ritmi di lavoro. Sui rischi psicosociali connessi all'*algorithmic management*, v. A. Cefaliello - P. V. Moore - R. Donoghue, *Making algorithmic management safe and healthy for workers*, *European Labour Law Journal*, 2023, n. 14, 192 ss.

⁽²⁷⁾ C. Maslach - M. Leiter, *Burnout at Work*, 2021.

⁽²⁸⁾ M.T. Carinci - S. Giudici - P. Perri, *Obblighi di informazione e sistemi decisionali e di monitoraggio automatizzati (art. 1-bis “Decreto Trasparenza”): quali forme di controllo per i poteri datoriali algoritmici?*, *Labor*, 2023,

Tali effetti non investono esclusivamente la dimensione dell'eguaglianza, ma si estendono anche al piano della tutela della salute. L'attribuzione automatizzata di valutazioni negative, i fenomeni di marginalizzazione algoritmica e l'impossibilità di ricostruire le logiche decisionali costituiscono infatti, a tutti gli effetti, fattori di *stress* lavoro-correlato e rischi psicosociali ⁽²⁹⁾. D'altra parte, l'utilizzo di sistemi non spiegabili può determinare «insicurezza professionale, pressione conformativa e perdita di controllo sul proprio lavoro» ⁽³⁰⁾.

Il co. 3, allora, assume un rilievo prevenzionistico diretto: l'assenza di discriminazioni non è soltanto un limite giuridico esterno, ma un indice materiale di sicurezza organizzativa. Cosicché, nel solco di tale lettura, il divieto di discriminazione si traduce in un parametro di salute e sicurezza: un algoritmo che discrimina risulta, al contempo, “illecito” – in quanto in violazione di un divieto espresso – e “pericoloso”, poiché idoneo a generare un rischio psicosociale giuridicamente rilevante ai sensi dell'art. 28 del d.lgs. n. 81/2008 ⁽³¹⁾.

2.3. La logica “precauzionale” nell'art. 11 tra prevenzione e governo dell'incertezza

Vi è però un ulteriore profilo che emerge dalla lettura congiunta dei tre commi dell'art. 11 e che la dottrina, sinora, non sembra aver adeguatamente valorizzato.

Il rinvio all'art. 1-bis, d.lgs. n. 152/1997 introduce, infatti, nell'ambito della disciplina della sicurezza del lavoro, un paradigma regolativo fino ad oggi rimasto sullo sfondo: “la precauzione” ⁽³²⁾.

1, 11 ss. Sul tema della *spiegabilità limitata* dei sistemi complessi, cfr. U. Ruffolo, *Artificial Intelligence Act e AI generativa*, cit., 440 ss., che discute dell'incapacità dell'utente di ricostruire la logica interna di modelli che operano tramite rappresentazioni non umane (“da modelli di linguaggio a modelli di mondo”).

⁽²⁹⁾ Cfr. anche T. Treu, *Intelligenza Artificiale (IA): integrazione o sostituzione del lavoro*, cit., 8 ss. secondo cui «(...) L'alternanza tra uomo e macchina intelligente produce nuove forme di pressione organizzativa e di rischio per i lavoratori».

⁽³⁰⁾ M.T. Carinci - S. Giudici - P. Perri, *Obblighi di informazione e sistemi decisionali*, cit., 467 ss..

⁽³¹⁾ È stato rilevato che «la razionalità numerica dell'algoritmo può produrre “esiti opachi e non ripercorribili, idonei a incidere sulla posizione soggettiva del lavoratore”» in L. Barbieri, *Sorveglianza umana di un sistema di LA*, cit., 2668. Sull'emergere del tema delle *allucinazioni algoritmiche* come fattore di rischio, cfr. G. Franco, *Responsabilità del professionista per uso improprio dell'LA*, LG, 3, 2026, 214, che richiama il caso del Trib. Firenze (14.3.2025) relativo all'impiego di riferimenti giurisprudenziali inesistenti generati dall'IA.

⁽³²⁾ Il principio di precauzione designa un criterio giuridico di intervento anticipato, in virtù del quale l'assenza di certezza scientifica non può giustificare l'inerzia decisionale a fronte di rischi potenzialmente gravi o irreversibili. Elaborato e applicato originariamente nel diritto internazionale ed europeo dell'ambiente e consacrato dal Principio 15 della Dichiarazione di Rio de Janeiro (1992), esso è positivizzato nell'ordinamento dell'Unione dall'art. 191, par. 2, TFUE e ha progressivamente assunto una valenza trasversale come logica di governo del rischio in condizioni di incertezza scientifica, anche oltre il settore ambientale. Per un approfondimento del tema, anche sul piano giurisprudenziale, v. M. V. Lumetti, *Il principio di precauzione nella legislazione, nella giurisprudenza e nelle recenti sentenze della Corte costituzionale*, Giur. amm., 2004, 42 ss.; L. Marini, *Il principio di precauzione nel diritto internazionale e comunitario*, Cedam, 2004.

L'obbligo informativo rafforzato rende visibili, tracciabili e ricostruibili gli elementi essenziali del funzionamento del sistema proprio perché il rischio, nell'ambito dell'IA, non è ancora pienamente conoscibile. L'informazione non descrive un rischio già definito: crea le condizioni affinché un rischio incerto, emergente e dinamico possa essere scrutinato, valutato e governato.

Sappiamo che il d.lgs. n. 81/2008 impone al datore di lavoro di valutare “tutti” i rischi (*ex artt.* 17, 28 e 29), integrando nella propria analisi anche le caratteristiche tecniche, adattive e algoritmiche dei sistemi di IA. Ma la prevenzione, per poter operare appieno, presuppone un grado minimo di conoscibilità del rischio: esige cioè che il perimetro dei pericoli sia definibile, misurabile e anticipabile.

Qui la prevenzione deve essere anticipata dalla *precauzione*.

Il co. 2 dell'art. 11, mediante il rinvio all'art. 1-*bis*, fornisce infatti la “materia prima” della valutazione: un'informazione strutturata che rende l'incertezza visibile e valutabile. Il Reg. (UE) 2024/1689 traduce tale logica in obblighi di documentazione, monitoraggio e tracciabilità continua, che trasformano il rischio incerto in contenuto probatorio ricostruibile e controllabile nel tempo.

Il co. 3 del medesimo articolo eleva il livello di garanzia sostanziale: impone di dimostrare l'assenza di discriminazioni algoritmiche come requisito intrinseco della liceità dello strumento.

A fronte di questa impostazione, il D.M. 17 dicembre 2025, n. 180 integra questi elementi nel sistema prevenzionistico, pur includendo i rischi psicosociali e organizzativi tra le aree da governare attraverso procedure di monitoraggio e vigilanza dedicate. Le Linee guida adottate con il decreto, pur formalmente qualificate come atto di indirizzo, assumono in realtà una portata sistematica: esse richiedono alle imprese un approccio «consapevole, sicuro e responsabile», fondato sulla supervisione umana, sulla documentazione delle logiche algoritmiche, sulla mitigazione dei rischi e sull'adozione di un sistema di *governance* dinamico, espressamente concepito per essere adattato all'evoluzione tecnologica, dal momento che «le Linee guida (...) sono soggette ad aggiornamento continuo».

Il documento ministeriale – coerente con i principi dell'*AI Act* e della l. n. 132/2025 – riconosce che l'IA può generare «*bias*, effetti distorsivi, rischi psicosociali e forme di pressione organizzativa» e prescrive «misure di controllo e vigilanza continua» per intercettarli ⁽³³⁾.

Ne scaturisce un modello – forse non consapevolmente ricercato dal legislatore – in cui prevenzione e precauzione si integrano con rara coerenza: la prima governa il rischio conoscibile, la seconda presidia l'incertezza.

⁽³³⁾ Per ulteriori elementi del D.M. 17 dicembre 2025, n. 180 che rafforzano la logica precauzionale, si vedano le seguenti previsioni: «Le Linee guida [...] sono soggette ad aggiornamento continuo da parte dell'Osservatorio sull'adozione dei sistemi di IA», a conferma del carattere dinamico dell'approccio regolativo; «Le decisioni che incidono su assunzioni, valutazioni, progressioni di carriera o licenziamenti devono prevedere supervisione umana, trasparenza e possibilità di contestazione», che ribadisce la necessità di un controllo umano costante nei processi decisionali algoritmici.

L'IA costringe il sistema prevenzionistico ad incorporare la logica della precauzione: la conoscenza non precede il rischio, ma viene prodotta per renderlo valutabile. L'informazione diventa condizione per trasformare l'incertezza in un contenuto probatorio governabile. E poiché l'IA introduce fisiologicamente incertezza, la precauzione diviene la condizione che rende praticabile la prevenzione, grazie all'informazione e, quindi, alla tracciabilità.

3. L'art. 12 e la *governance* conoscitiva dell'algoritmo: natura, ruolo e limiti dell'Osservatorio

L'atteggiamento del legislatore domestico nel maneggiare una questione ad alto tasso di incertezza – quale l'impiego dell'IA nei processi lavorativi – si traduce nell'art. 12 della legge in commento che istituisce l'Osservatorio sull'adozione dell'IA nel mondo del lavoro con l'obiettivo dichiarato di «massimizzare i benefici e contenere i rischi».

Si tratta, tuttavia, di un organismo a vocazione generalista, non specificamente deputato alla tutela della salute e sicurezza sul lavoro, né concepito come presidio specialistico dei rischi psicosociali e organizzativi connessi all'impiego dell'IA.

Proprio per tale ragione, ci si confronta con un soggetto la cui funzione resta solo implicitamente delineata: colmare il tradizionale vuoto informativo che caratterizza questa tipologia di rischi e trasformare la conoscenza in capacità di intervento, attraverso la produzione di dati comparabili, l'elaborazione di indicatori confutabili e l'adozione di meccanismi di aggiornamento periodico idonei a intercettare l'evoluzione tecnologica e organizzativa.

Questa impostazione è confermata dal documento ministeriale che accompagna l'avvio dei lavori, il quale qualifica l'Osservatorio come strumento destinato a «offrire una base conoscitiva ampia e articolata sui principali ambiti di incidenza dell'IA» e a costituire «un punto di partenza per un percorso di approfondimento e collaborazione» volto allo sviluppo di analisi e orientamenti aggiornabili nel tempo ⁽³⁴⁾.

Ebbene, tali prescrizioni si fondano su un presupposto comune: che il rischio connesso all'impiego dei sistemi di IA sia previamente individuabile e conoscibile attraverso strumenti istruttori adeguati. È questo assetto, nel quale la conoscenza del rischio costituisce il fulcro della *governance*, a rivelare i suoi limiti nei contesti lavorativi caratterizzati da elevata complessità organizzativa e da effetti indiretti e cumulativi sulla salute e sicurezza.

È a questo livello, infatti, che emergono ulteriori profili critici, che si aggiungono a quelli già rilevati e che devono essere letti alla luce del rapporto tra prevenzione e

⁽³⁴⁾ Ministero del Lavoro e delle Politiche Sociali, *Verso l'Osservatorio sull'adozione di sistemi di intelligenza artificiale nel mondo del lavoro. Raccolta di contributi a supporto dell'avvio dei lavori dell'Osservatorio*, febbraio 2026.

precauzione, specie quando il rischio non sia immediatamente misurabile né pienamente spiegabile in via *ex ante*, come accade nei casi di impiego dell'intelligenza artificiale.

Il primo profilo critico che emerge riguarda l'effettività della supervisione umana: affinché essa sia coerente con le esigenze di tutela della salute e sicurezza sul lavoro, deve consentire non solo una valutazione reale degli esiti del sistema, ma anche la correzione tempestiva di decisioni distorte e, ove necessario, la sospensione o l'interruzione del funzionamento algoritmico. Tale effettività, tuttavia, presuppone la disponibilità di un patrimonio conoscitivo adeguato, idoneo a rendere il rischio intelligibile e valutabile sul piano organizzativo e prevenzionistico.

In questa direzione, il ruolo attribuito all'Osservatorio assume un rilievo centrale ma al tempo stesso problematico: la supervisione umana rischia infatti di restare meramente formale laddove la conoscenza dei rischi – in particolare quelli psicosociali e di natura cumulativa – non sia sufficientemente strutturata, costantemente aggiornata e concretamente traducibile in criteri operativi di intervento.

Un secondo profilo critico attiene alla gestione dei dati che l'Osservatorio è chiamato a raccogliere. La funzione conoscitiva comporta l'acquisizione di informazioni su modelli, incidenti, criticità emergenti; tuttavia, questa attività non può trasformarsi in un nuovo punto di vulnerabilità. Il Reg. (UE) 2016/679 impone una protezione dei dati fin dalla progettazione: minimizzazione, controllo degli accessi, anonimizzazione adeguata e valutazioni d'impatto ove necessarie. La trasparenza dell'IA, in questa prospettiva, deve concretizzarsi in evidenze documentali strutturate – registri ordinati, versioni tracciabili, controlli verificabili – idonee a sostenere sia la dimensione precauzionale sia quella prevenzionistica.

Un terzo profilo riguarda il metodo di lavoro dell'Osservatorio.

La composizione multilivello dell'Osservatorio, pur garantendo pluralismo e confronto tra attori sociali, conferma la sua vocazione prevalentemente conoscitiva e consultiva, lasciando aperto il problema della trasformazione di tale patrimonio informativo in criteri operativi effettivamente utilizzabili nel sistema prevenzionistico della salute e sicurezza sul lavoro.

La presenza delle rappresentanze sindacali costituisce certamente una base solida, ma non sufficiente. Il documento ministeriale sottolinea, infatti, che i contributi raccolti riflettono «approcci, sensibilità e punti di vista diversi» e che l'obiettivo dichiarato è quello di costruire strumenti «sempre più puntuali a supporto di un'adozione consapevole, trasparente ed efficace dell'IA».

Perché tale finalità possa tradursi in tutela effettiva, la partecipazione degli attori coinvolti deve tuttavia assumere una dimensione propriamente operativa, orientata alla verifica della qualità dei dati, alla valutazione dei modelli, all'interpretazione dei segnali di rischio e alla definizione di metodologie condivise. Solo in questa prospettiva l'Osservatorio può evolvere in una vera e propria “infrastruttura pubblica di conoscenza”, divenendo uno spazio nel quale l'incertezza viene progressivamente resa conoscibile – e dunque governabile – attraverso la produzione di indicatori affidabili e concretamente

utilizzabili dai soggetti che presidiano il sistema prevenzionistico, quali ispettorati, parti sociali e giudici.

Anche la dimensione economica e organizzativa rileva in termini problematici. Dai contributi ministeriali emergono, infatti, ritardi diffusi nell'adozione dell'IA, significativi divari territoriali e dimensionali, nonché le difficoltà delle piccole e medie imprese, riconducibili a costi elevati, carenze di competenze e persistente incertezza regolativa. In tale contesto, la funzione dell'Osservatorio consiste anche nel distinguere fenomeni diversi – automazione, creazione di nuove mansioni, tecnologie che potenziano il lavoro umano o principalmente il capitale – perché solo una corretta qualificazione consente una valutazione realistica dei rischi e una definizione proporzionata delle priorità di prevenzione.

Da questo quadro emerge, tuttavia, un ulteriore elemento di tensione: il legislatore non ha attribuito all'Osservatorio né poteri regolativi, né compiti di verifica tecnica, né obblighi di raccordo stabile con i presidi prevenzionistici previsti dal d.lgs. n. 81/2008. Ne consegue che l'Osservatorio si configura come uno strumento di coordinamento e produzione conoscitiva, fondato su una cooperazione volontaria tra soggetti pubblici e privati e caratterizzato da una struttura conoscitiva prevalentemente descrittiva, più che prescrittiva.

Ciò disvela che prevenzione e precauzione, pur entrambe richiamate implicitamente dal legislatore, non si sovrappongono né confluiscono in un modello unitario: costituiscono piuttosto due dimensioni che l'art. 12 tenta di avvicinare, senza tuttavia riuscire a integrarle compiutamente.

L'Osservatorio può contribuire a ridurre questa distanza trasformando l'incertezza in dati strutturati, ma – come rilevano gli stessi materiali ministeriali, costruiti come «raccolta strutturata di contributi eterogenei» che rappresentano «punti di vista diversi» e non un indirizzo unitario – la funzione conoscitiva è ancora embrionale, disomogenea e priva di un metodo vincolante ⁽³⁵⁾. La sua capacità di incidere sulla tutela della salute e sicurezza dipenderà quindi non da una previsione normativa immediatamente cogente, ma dalla concreta evoluzione del metodo di lavoro, dalla qualità dei dati raccolti e dalla capacità di tradurre l'eterogeneità dei contributi in orientamenti utilizzabili.

Per tali ragioni, l'art. 12 non può essere letto come un punto di arrivo né come un presidio sufficiente del rischio derivante dall'IA. Rappresenta piuttosto un tentativo iniziale e ancora imperfetto di costruire una base conoscitiva comune, indispensabile per rendere praticabile la prevenzione ma, allo stato, non idonea a colmare il *deficit* informativo che grava su rischi organizzativi, psicosociali e discriminatori. La sua effettività, pertanto, resta un obiettivo da costruire, non un risultato già conseguito ⁽³⁶⁾.

⁽³⁵⁾ Per una lettura sistematica del ruolo degli Osservatori ministeriali e della loro funzione non regolatoria, v. F. Costantini, *La L. 132/2025 sull'intelligenza artificiale*, cit., 1096, che sottolinea la natura eminentemente informativa e la dipendenza dagli atti attuativi.

⁽³⁶⁾ Una parte delle soluzioni richiamate – in particolare l'esigenza di un raccordo strutturale tra funzione conoscitiva, prevenzione del rischio e logica precauzionale nell'impiego dei sistemi di IA – era già

4. L'art. 13 e i limiti professionali dell'ausilio algoritmico

L'art. 13, l. n. 132/2025 disciplina l'utilizzo dei sistemi di IA nell'esercizio delle professioni intellettuali. La norma non si rivolge agli "operatori della salute e sicurezza" in via esclusiva, ma a tutti i professionisti, sia organizzati in ordini sia appartenenti alle professioni non regolamentate, come conferma espressamente la prassi interpretativa: l'art. 13 «si applica a tutte le professioni intellettuali, sia organizzate in ordini che non» ⁽³⁷⁾.

Tuttavia, nel presente contributo l'attenzione è rivolta in modo particolare agli operatori tecnici della prevenzione e alle professioni che intervengono nella gestione dei rischi, poiché l'impatto dell'IA sulle loro responsabilità presenta tratti specifici che non si riscontrano con la medesima intensità negli altri ambiti professionali disciplinati dall'art. 13. Benché la norma si applichi indistintamente «a tutte le professioni intellettuali, sia organizzate in ordini che non», nel settore della salute e sicurezza sul lavoro l'interazione fra valutazione tecnico-professionale, supervisione umana e trasparenza sull'uso dell'IA assume un rilievo peculiare, in quanto incide direttamente sulla validità delle analisi dei rischi, sulla tracciabilità delle scelte e sulla imputabilità delle responsabilità.

La disposizione impone la centralità del contributo umano: l'IA è ammessa unicamente quale strumento ausiliario, in coerenza con il principio per cui «l'utilizzo di sistemi di intelligenza artificiale nelle professioni intellettuali è finalizzato al solo esercizio delle attività strumentali e di supporto all'attività professionale» con «prevalenza del lavoro intellettuale» del professionista.

Ne deriva che la trasparenza informativa assume valore operativo: il professionista è tenuto a comunicare «con linguaggio chiaro, semplice ed esaustivo» al cliente o al destinatario tecnico della prestazione l'uso dell'intelligenza artificiale e le sue funzioni nello svolgimento dell'incarico.

Per gli operatori della salute e sicurezza – destinatari di obblighi di controllo, verifica e giudizio tecnico – questa prescrizione si traduce nella necessità di rendere ricostruibile

stata delineata, sia pure in forma embrionale, in altra sede: v. M. Squeglia, *Obiettivi, strumenti e obiettivi dell'intelligenza artificiale*, cit., spec. 131 ss.

⁽³⁷⁾ Sul contenuto concreto degli obblighi informativi introdotti dall'art. 13, l. n. 132/2025, si veda R. Braga, *AI e professionisti: un nuovo modello per gli obblighi informativi verso i clienti*, *Pratica Fiscale e Professionale*, n. 45/2025, 33-36, dalla quale emerge che l'IA può essere utilizzata dai professionisti soltanto per attività strumentali e di supporto, mentre l'attività intellettuale deve restare «in ogni caso prevalente». L'A. segnala inoltre che: l'obbligo informativo riguarda tutti i professionisti, indipendentemente dall'iscrizione ad ordini; – l'informativa deve indicare in modo chiaro, semplice ed esaustivo il tipo di sistemi di IA impiegati, le finalità del loro utilizzo, i rischi, i margini di errore e le modalità di verifica e validazione degli output; – il Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Contabili ha diffuso un esempio di clausola contrattuale da inserire negli incarichi professionali (Comunicato stampa 9 ottobre 2025), mentre Confprofessioni e l'Associazione Nazionale Forense hanno predisposto un modello completo di informativa per i clienti, che evidenzia la natura meramente strumentale dell'IA e la necessità di un controllo umano costante; – la legge non prevede sanzioni amministrative specifiche, ma il mancato adempimento dell'obbligo informativo integra una violazione dei doveri deontologici, con conseguente responsabilità disciplinare.

l'intero processo professionale in cui l'IA è stata utilizzata: quali dati sono stati impiegati, quale versione del modello, quali parametri sono stati applicati, quali alternative valutate⁽³⁸⁾. In altre parole, la trasparenza deve diventare tracciabilità tecnica, affinché la responsabilità rimanga pienamente umana e imputabile⁽³⁹⁾. La dottrina ha osservato che l'art.13 mira a garantire che la responsabilità «rimanga interamente a carico del professionista» anche in caso di uso di strumenti avanzati di IA, richiedendo una valutazione critica ed una supervisione effettiva sull'*output* algoritmico.

Rimangono, tuttavia, alcune criticità interpretative e applicative che la norma non risolve in modo definitivo.

Una prima criticità attiene all'ambito oggettivo dell'obbligo informativo: se è vero che la legge richiede che l'informativa sia resa in forma "esaustiva", resta tuttavia indeterminato quali categorie di sistemi debbano esserne ricomprese, né quale sia la soglia minima di completezza idonea a garantire la piena conformità sul piano legale. In particolare, non risulta chiarito se i sistemi di IA per uso generale – quali, ad es., *ChatGPT*, *Claude* o strumenti analoghi – debbano ritenersi automaticamente inclusi nell'obbligo informativo, ovvero se la loro rilevanza dipenda dalla concreta modalità con cui essi intervengono nel processo professionale e incidono sull'elaborazione della prestazione.

Un secondo profilo critico emerge in relazione all'impatto del pluralismo deontologico. Le professioni tecniche, incluse quelle operanti nell'ambito della prevenzione, risultano disciplinate da regimi di autoregolazione tra loro differenziati, che possono presentare divergenze significative quanto alle forme di trasparenza richieste, al livello di tracciabilità dell'attività svolta e alla qualità dell'informazione dovuta al destinatario della prestazione. In assenza di linee guida comuni e di modelli standardizzati, tale eterogeneità rischia di tradursi in una frammentazione degli standard applicativi, con conseguenti disallineamenti sia nella qualificazione della responsabilità sia nella prova dell'effettivo adempimento degli obblighi informativi⁽⁴⁰⁾.

⁽³⁸⁾ Cfr. G. Comandé, *La responsabilità civile dopo la L. n. 132/2025: per favore parliamone!*, *Danno e responsabilità*, 6, 2025, 672 ss.

⁽³⁹⁾ Sulla responsabilità del professionista che faccia uso imprudente di sistemi di IA, cfr. G. Franco, *Responsabilità del professionista*, cit., 214 ss., il quale analizza le prime applicazioni giurisprudenziali della materia. Allo stato, tali pronunce riguardano prevalentemente l'attività dell'avvocato, senza che si registrino ancora pronunce idonee a estendere tali approdi ad altre figure professionali al di fuori dell'ambito forense. In particolare, si tratta delle decisioni Trib. Torino, 16 settembre 2025, n. 2120, che ha condannato la parte attrice ai sensi dell'art. 96, co. 3, c.p.c., per avere proposto un ricorso redatto "con il supporto dell'intelligenza artificiale", caratterizzato da un "coacervo di citazioni normative e giurisprudenziali astratte, prive di ordine logico e in larga parte inconferenti, senza alcuna verifica della loro attendibilità"; nonché Trib. Latina, 23 settembre 2025, conforme, che ha parimenti ravvisato un uso distorto dell'IA nella redazione di atti giudiziari privi di riscontri fattuali, evidenziando come tale condotta integri una forma di abuso del processo e giustifichi la condanna ex art. 96 c.p.c.

⁽⁴⁰⁾ Sul piano deontologico-operativo, si registrano già modelli non del tutto omogenei: il Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Contabili (Comunicato stampa 9 ottobre 2025) ha predisposto una "clausola contrattuale tipo" da inserire nel mandato professionale, centrata sulla natura meramente ausiliaria dell'IA, sulla permanenza della responsabilità in capo al professionista e sul diritto di opposizione del cliente; Confprofessioni e Associazione Nazionale Forense, invece, hanno predisposto

La norma si limita a enunciare tale principio in termini generali, senza però introdurre criteri operativi idonei a definire in quale misura l'IA possa incidere sui processi decisionali, né a individuare il livello di intervento umano necessario affinché la prestazione possa dirsi effettivamente riconducibile al professionista. Ne deriva una persistente indeterminatezza circa il “quanto” e il “dove” il giudizio umano debba concretamente prevalere, con la conseguente emersione di aree di incertezza sotto il profilo della responsabilità civile e disciplinare. Per evitare tali distorsioni applicative, appare necessario prevedere criteri omogenei, riconducibili almeno a: soglie uniformi di informazione, volte a individuare i contenuti che devono essere sempre comunicati; parametri coerenti di tracciabilità, funzionali a documentare le modalità di impiego dell'IA nell'attività professionale; criteri condivisi di spiegabilità, destinati a delimitare il grado di intellegibilità richiesto in relazione all'uso dei sistemi algoritmici.

In questo quadro si inserisce, infine, un ulteriore elemento problematico rappresentato dall'assenza di un apparato sanzionatorio specifico per il caso di inadempimento dell'obbligo di informare il cliente circa l'impiego di sistemi di IA. La scelta legislativa di rinviare ai doveri deontologici già esistenti, sebbene coerente con la natura fiduciaria della prestazione professionale, rischia di attenuare l'effettività della tutela, affidando la cogenza della norma a meccanismi settoriali ed eterogenei, non sempre idonei a garantire uniformità applicativa e certezza delle conseguenze sul piano della responsabilità.

Alla luce di tali criticità, le soglie minime di informazione, tracciabilità e spiegabilità devono essere calibrate in modo coerente con il quadro regolatorio vigente – segnatamente il d.lgs. n. 81/2008, il Reg. (UE) 2016/679 e il Reg. (UE) 2024/1689 – poiché solo all'interno di un simile coordinamento sistematico la responsabilità può rimanere imputabile, effettivamente controllabile e funzionalmente allineata alle condizioni di liceità del trattamento tecnico-professionale.

5. Conclusioni: dalla dichiarazione alla prova

Il percorso complessivo delineato dagli artt. 11-13, l. n. 132/2025, restituisce un quadro che – sia pure non privo di distonie – non si esaurisce nell'introduzione di nuovi obblighi informativi o in un insieme di adempimenti meramente formali.

Esso pare richiedere, piuttosto, una diversa concezione della responsabilità, fondata non sull'enunciazione delle intenzioni, ma sulla “ricostruibilità” delle scelte che hanno guidato l'impiego dei sistemi di IA, in coerenza con i vincoli e le finalità poste dal diritto dell'Unione europea.

un'informativa più analitica, che richiede l'indicazione delle tipologie di IA utilizzate, della loro collocazione (UE/extra-UE), dei rischi e delle prassi di verifica e validazione degli output, con acquisizione formale della presa d'atto del cliente. Tali differenze confermano il rischio di *standard* applicativi disallineati in assenza di linee guida comuni.

È forse proprio in questo aspetto che può essere colto il valore delle norme in commento: la trasparenza non si configura come un attributo meramente dichiarativo, né come una formula di stile priva di contenuto operativo, ma come un vero e proprio criterio probatorio, idoneo a misurare la liceità del sistema, la correttezza del processo decisionale e la verificabilità dell'intervento umano che lo accompagna.

La premessa metodologica è inscritta nell'art. 11, co. 1, che orienta l'impiego dell'IA nei rapporti di lavoro al miglioramento delle condizioni di lavoro, alla tutela dell'integrità psicofisica, all'accrescimento della qualità delle prestazioni e all'aumento della produttività, traducendo sul piano legislativo il bilanciamento tra art. 32 e art. 41 Cost. Da questa biforcazione discende una conseguenza essenziale: la qualificazione della finalità concreta per cui un sistema viene adottato determina non solo la sua liceità, ma l'intero percorso regolativo che deve seguirne l'introduzione.

È qui che si colloca il nodo critico che la normativa, da sola, non è in grado di risolvere: come garantire che la finalità dichiarata dal datore di lavoro corrisponda a quella effettivamente perseguita dal sistema.

Il mero affidamento sull'autodichiarazione datoriale si rivela inadeguato, soprattutto in presenza di sistemi polifunzionali o *general purpose*, nei quali l'impatto sull'organizzazione del lavoro non è predeterminato dalla tecnologia in astratto, ma dipende in modo decisivo dalle modalità concrete di integrazione e di utilizzo nel contesto produttivo ⁽⁴¹⁾. Per questa ragione, il quadro europeo insiste su documentazione, tracciati di *audit*, storia delle versioni e ricostruzione dell'intero ciclo di vita dei sistemi: strumenti che consentono di evitare che la trasparenza si riduca a una formula meramente dichiarativa e priva di verificabilità.

L'esperienza maturata nel contrasto al *greenwashing* ⁽⁴²⁾ – inteso come la prassi di attribuire caratteri di sostenibilità, in particolare quella ambientale, non supportati da evidenze verificabili – offre, tuttavia, un parallelo utile ma non meramente suggestivo: quando le dichiarazioni di sostenibilità non sono accompagnate da meccanismi di verifica indipendente, esse degenerano in strumenti comunicativi privi di fondamento fattuale.

Allo stesso modo, in materia di IA, termini come “affidabilità”, “sicurezza” o “trasparenza” rischiano di diventare espressioni retoriche, se non sono sostenute da procedure verificabili.

In questa prospettiva, una risposta coerente con l'impianto dell'art. 11 potrebbe consistere nell'introduzione di una certificazione tecnico-indipendente delle finalità del sistema.

Un organismo terzo, accreditato secondo criteri omogenei e controllabili, potrebbe infatti attestare se un'applicazione di IA è adottata per finalità riconducibili all'art. 32 Cost., caso nel quale la certificazione costituirebbe il presupposto per l'attivazione del percorso

⁽⁴¹⁾ P. Pascucci, *Legal aspects of health and safety at work in Italy: recent trends and perspectives*, *AmbienteDiritto.it*, XXIV, 1, 2024.

⁽⁴²⁾ Cfr. S. Bogojević, *Regulating Greenwashing: Where Market Rules Meet Climate Action*, *Journal of Environmental Law*, vol. 37, 2, 2025, 169 ss. ove si sottolinea che, in assenza di sistemi di verifica indipendente, la trasparenza delle dichiarazioni ambientali resta priva di reale capacità regolativa.

informativo rafforzato previsto dall'art. 1-*bis*, d.lgs. n. 152/1997 e per la conseguente integrazione nella valutazione dei rischi *ex art.* 28, d.lgs. n. 81/2008.

Se invece l'attività fosse destinata a finalità organizzativo-produttive riconducibili all'art. 41 Cost., l'introduzione del sistema richiederebbe, prima ancora di qualunque valutazione prevenzionistica, il rispetto della procedura autorizzatoria o negoziale *ex art.* 4 St. lav., trattandosi di un sistema potenzialmente idoneo a incidere sull'organizzazione, sulle modalità di esecuzione della prestazione o sui controlli a distanza.

A nostro avviso, già la sola previsione dell'esistenza di un organismo terzo incaricato di tale certificazione avrebbe un effetto dissuasivo significativo: il datore di lavoro difficilmente si rivolgerebbe ad un simile organismo quando fosse consapevole, già *ex ante*, che il sistema è adottato per finalità diverse o ulteriori rispetto a quelle indicate dal legislatore e, in particolare, differenti da quella di «*tutelare l'integrità psicofisica dei lavoratori*» di cui all'art. 11, co. 1 ⁽⁴³⁾.

Solo dopo questa qualificazione preventiva – e non sulla base di una ricostruzione *ex post* – il sistema potrebbe essere correttamente collocato nella valutazione dei rischi o nei processi organizzativi.

La certificazione, in questo senso, non sostituisce la prevenzione: ne costituisce la condizione di intelligibilità, poiché rende verificabile la finalità concretamente perseguita dal sistema e consente una applicazione non arbitraria delle norme richiamate.

Sotto questo profilo, l'Osservatorio istituito dall'art. 12, pur nella sua natura generalista e priva di funzioni regolative, potrebbe svolgere un ruolo essenziale non sul piano delle autorizzazioni, ma come struttura conoscitiva. Esso potrebbe contribuire a individuare le aree di “rischio precauzionale”, trasformando dati eterogenei in indicatori verificabili e confutabili, così da riportare l'incertezza prodotta dall'IA entro i confini ordinari della prevenzione.

A ben vedere, il diritto vigente contiene già gli strumenti necessari per colmare lo scarto tra dichiarazioni e prove: documentazione, tracciabilità, supervisione umana effettiva, percorsi autorizzatori, valutazione dei rischi, obblighi informativi rafforzati.

La sfida risiede, piuttosto, nel tradurre tale apparato in pratiche verificabili, idonee ad assicurare che gli obblighi normativi non rimangano confinati ad una dimensione meramente dichiarativa ma trovino effettiva attuazione nei processi tecnici e organizzativi

⁽⁴³⁾ La nostra proposta non determina alcuna sovrapposizione con gli organismi di valutazione della conformità previsti dall'*AI Act*, disciplinati dagli artt. 33 e segg., Reg. (UE) 2024/1689. Tali organismi svolgono una funzione tecnica specifica: *audit* algoritmici e verifiche ingegneristiche su sicurezza, robustezza, qualità dei dati, gestione del rischio e *governance* del ciclo di vita dei sistemi di IA, secondo le norme europee e gli standard ISO/IEC 42001. L'organismo indipendente delineato nel presente contributo ha invece natura sì tecnica, ma orientata alla qualificazione tecnico-giuridica della finalità di impiego del sistema nel contesto lavorativo: è chiamato a verificare, attraverso analisi funzionali e valutazioni tecniche, se l'uso dell'IA sia riconducibile all'art. 32 Cost. o all'art. 41 Cost., così da individuare *ex ante* il corretto percorso regolativo tra d.lgs. n. 152/1997, art. 4 St. lav. e d. lgs. n. 81/2008. Le due funzioni sono quindi complementari, ma non sovrapponibili: la prima attiene alla conformità tecnica del modello, la seconda alla qualificazione tecnico-giuridica della sua finalità effettiva.

La documentazione delle versioni del modello, dei *dataset* e dei parametri di *input* utilizzati assume così la funzione di una vera e propria “prova della finalità”: non è sufficiente dichiarare lo scopo per cui il sistema è adottato, ma è necessario dimostrare che le scelte tecniche operate – in termini di architettura, addestramento, configurazione e modalità d’uso – siano coerenti e funzionalmente orientate al perseguimento di quella finalità dichiarata ⁽⁴⁴⁾.

Ciò presuppone l’impiego della documentazione quale strumento probatorio, un ruolo attivo dell’Osservatorio nella produzione di conoscenza tecnica pubblica ed un impegno delle professioni – in particolare, di quelle della prevenzione e della sorveglianza sanitaria – nell’assunzione della trasparenza come *standard* etico ed operativo e non come mero adempimento formale ⁽⁴⁵⁾.

⁽⁴⁴⁾ Sulla necessità di distinguere tra affidabilità dichiarata e affidabilità dimostrabile, cfr. U. Ruffolo, *Artificial Intelligence Act e AI generativa*, cit., 442 ss. ove si evidenzia che i concetti di “affidabilità” e “sicurezza” rischiano, in assenza di verifiche, di ridursi a mere formule retoriche.

⁽⁴⁵⁾ V. anche L. Barbieri, *Sorveglianza umana di un sistema di LA*, cit., 2669.

Bibliografia

- Alvino I., *I nuovi limiti al controllo a distanza dell'attività dei lavoratori nell'intersezione fra le regole dello Statuto dei lavoratori e quelle del Codice della privacy*, in LLI, vol. 2, 1, 2016, 35 ss.
- Barbieri L., *Sistemi di LA nel mondo del lavoro: Linee guida ministeriali ed elementi di raccordo con la normativa vigente*, in DPL, 28, 2025, 1710 ss.
- Barbieri L., *Sorveglianza umana di un sistema di LA ad alto rischio*, in DPL, 46, 2025, 2666 ss.
- Bavaro V., *Lavoro tramite piattaforme digitali e subordinazione algoritmica*, in ADL, 2021, 211 ss.
- Bogojević S., *Regulating Greenwashing: Where Market Rules Meet Climate Action*, in *Journal of Environmental Law*, vol. 37, 2, 2025, 169 ss.
- Braga R., *AI e professionisti: un nuovo modello per gli obblighi informativi verso i clienti*, in *Pratica Fiscale e Professionale*, n. 45/2025, 33 ss.
- Cairolì S., *Intelligenza artificiale e sicurezza sul lavoro: uno sguardo oltre la siepe*, in DSL, 2, 2024, 26 ss.
- Carinci M. T. - Giudici S. - Perri P., *Obblighi di informazione e sistemi decisionali e di monitoraggio automatizzati (art. 1-bis "Decreto Trasparenza"): quali forme di controllo per i poteri datoriali algoritmici?*, in *Labor*, 2023, 1, 7 ss.
- Cefaliello A. - Moore P. V. - Donoghue R., *Making algorithmic management safe and healthy for workers*, in *European Labour Law Journal*, 2023, n. 14, 192 ss.
- Ciucciovino S., *Intelligenza artificiale e diritto del lavoro: problemi e prospettive - Risorse umane e intelligenza artificiale alla luce del regolamento (UE) 2024/1689, tra norme legali, etica e codici di condotta*, in DRI, 3, 2024, 573 ss.
- Comandé G., *La responsabilità civile dopo la L. n. 132/2025: per favore parliamone!*, in *Danno e responsabilità*, 6, 2025, 672 ss.
- Costantini F., *La L. n. 132/2025 sull'intelligenza artificiale: una prima lettura e alcune indicazioni in materia di diritto del lavoro*, in LG, 12, 2025, 1093 ss.
- Cristianini N., *Machina sapiens. L'algoritmo che ci ha rubato il segreto della conoscenza*, Il Mulino, 2024.
- De Petris, P., *La questione della trasparenza algoritmica nei rapporti di lavoro interessati dal management automatizzato: una lettura integrata*, in MGL, 4, 2024, 781 ss.
- Elmo M. G., *Sistemi LA e rischi per la salute e la sicurezza dei lavoratori: riflessioni a margine della regolamentazione europea*, in *AmbienteDiritto.it*, Anno XXIV, fasc. 4, 2024, 1 ss.
- Faini F., *Intelligenza artificiale e regolazione giuridica: il ruolo del diritto nel rapporto tra uomo e macchine*, in *federalismi.it*, 2023, n. 2, 1 ss.
- Faioli M., *Adeguatezza ex art. 2086 c.c. e obbligo di introdurre tecnologia avanzata per mitigare i rischi da lavoro*, in *federalismi.it*, 2025, fasc. 6, 143 ss.
- Faioli M., *Robot Labor Law. Linee di ricerca per una nuova branca del diritto del lavoro e in vista della sessione sull'intelligenza artificiale del G7 del 2024*, in *Federalismi.it*, 2024, n. 8, 182 ss.
- Floridi L., *AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models*, in *Philosophy & Technology*, 2023, 36(15).
- Franco G., *Responsabilità del professionista per uso improprio dell'IA*, in LG, 3, 2026, 213 ss.
- Frascheri C., *Gestire il cambiamento*, Edizioni Lavoro, 2023.
- Ghera E., *L'articolo 39 della Costituzione e il contratto collettivo*, in WP C.S.D.L.E. "Massimo D'Antona", n. 202/2014.
- Gragnoletti E., *I controlli sul lavoratore e i cosiddetti sistemi di intelligenza artificiale*, in LDE, 2024, n. 3, 1 ss.
- Ingrao A., *Intelligenza artificiale nei rinnovi dei contratti collettivi nazionali di categoria. Una prima roadmap*, in LDE, 27 ottobre 2025.
- Lazzari C., *Lavoro senza luogo fisso, dematerializzazione degli spazi, salute e sicurezza*, in LLI, 2023, n. 1, 22 ss.
- Loi P., *Il rischio proporzionato nella proposta di regolamento sull'IA*, in *federalismi.it*, 2023, 4, 241 ss.
- Lumetti M. V., *Il principio di precauzione nella legislazione, nella giurisprudenza e nelle recenti sentenze della Corte costituzionale*, in *Giur. amm.*, 2004, 42 ss.

- Maio V., *Il diritto del lavoro e le nuove sfide della rivoluzione robotica*, in *ADL*, 2018, n. 6, 1414 ss.
- Marassi S., *Intelligenza artificiale e sicurezza sul lavoro*, in M. Biasi (a cura di), *Diritto del lavoro e intelligenza artificiale*, Giuffrè, 2024, 207 ss.
- Marini L., *Il principio di precauzione nel diritto internazionale e comunitario*, Cedam, 2004.
- Maslach C. - Leiter M., *Burnout at Work*, 2021.
- Novelli C., *L'Artificial Intelligence Act europeo: alcune questioni di implementazione*, in *federalismi.it*, 2024, 2, 96 ss.
- Pascucci P. - Lazzari C., *Sistemi di IA, salute e sicurezza sul lavoro: una sfida al modello di prevenzione aziendale, fra responsabilità e opportunità* in *DSL*, 4, 2024, 587 ss.
- Pascucci P., *Le nuove coordinate del sistema prevenzionistico*, in *DSL*, 2023, 2, I, 48 ss.
- Pascucci P., *Legal aspects of health and safety at work in Italy: recent trends and perspectives* in *AmbienteDiritto.it*, XXIV, 1, 2024, 1 ss.
- Pasquarella V., *I principali garanti della sicurezza nell'era delle nuove trasformazioni del lavoro*, in *VTDL*, n. 4, 2023, 890 ss.
- Peruzzi M., *Sistemi automatizzati e tutela della salute e sicurezza sul lavoro*, in *DSL*, 2024, n. 2, 78 ss.
- Romeo C., *L'era degli algoritmi e la sua incidenza nell'ambito della certezza del diritto: un connubio sospetto*, in *LG*, 2024, n. 1, 5 ss.
- Rosin A. - Parviainen H., *Improving the transparency of algorithmic management in platform work*, in *European Labour Law Journal*, 2025, vol. 16(4), 472 ss.
- Rotondi F. - Tursi A., *I controlli sull'attività dei lavoratori alla prova dell'intelligenza artificiale*, in *LG*, 2, 2025, 113 ss.
- Ruffolo U., *Artificial Intelligence Act e AI generativa*, in *GI*, 2025, 2, 438 ss.
- Ruffolo U., *Per i fondamenti di un diritto della robotica self-learning: dalla machinery produttiva all'auto driverless: verso una «responsabilità da algoritmo»?* , in *Intelligenza artificiale e responsabilità*, Giuffrè, 2017, 1 ss.
- Squeglia M., *Obiettivi, strumenti e obiettivi dell'intelligenza artificiale nella gestione della sicurezza del lavoro*, in *DSL*, 1, 2025, 114 ss.
- Tomaszewska M., *Transparency and Algorithmic Oversight in the Digital Workplace: Legal Challenges under Directive 2024/2831*, in *Studia z Zakresu Prawa Pracy i Polityki Społecznej*, 2025, vol. 32, n. 3, 187 ss.
- Treu T., *Intelligenza Artificiale (IA): integrazione o sostituzione del lavoro umano?*, in *CSDLE "Massimo D'Antona".IT*, n. 487/2024.
- Tullini P. (a cura di), *Controlli a distanza e tutela dei dati personali del lavoratore*, Giappichelli, 2017.
- Tullini P., *La Direttiva piattaforme e i diritti del lavoro digitale*, in *LLI*, 2022, 1, 1 ss.
- Yannakourou S., *Regulating Algorithmic Management of Work at EU Level: The Adequacy of the Platform Work Directive and the AI Act*, in *EU Digital Law in the AI Era*, Springer, 2025, 33 ss.
- Zappalà L., *Informatizzazione dei processi decisionali e diritto del lavoro: algoritmi, poteri datoriali e responsabilità*, in *CSDLE "Massimo D'Antona"*, 2021, n. 446.
- Zoli C., *Il controllo a distanza dell'attività dei lavoratori e la nuova struttura dell'articolo 4 della legge 300/1970*, in *VTDL*, 4, 2016, 635 ss.