Intelligenza Artificiale e Salute e Sicurezza sul Lavoro (SSL): la Sfida dell'Equilibrio

Scheda Rete Giuridica
Roberto Sammarchi, Avvocato specialista in diritto dell'informazione, della comunicazione digitale e della protezione dei dati personali.

Questa scheda esamina il delicato equilibrio tra l'uso dell'Intelligenza Artificiale (IA) per migliorare la Salute e Sicurezza sul Lavoro (SSL) e la navigazione nel complesso panorama normativo, in particolare il Regolamento UE sull'IA. Esplora il potenziale trasformativo dell'IA nella prevenzione degli incidenti sul lavoro e nella promozione del benessere dei lavoratori attraverso l'analisi predittiva, il monitoraggio in tempo reale e i sistemi di sicurezza automatizzati. Il documento sottolinea la necessità di affrontare considerazioni etiche, preoccupazioni sulla privacy dei dati e il potenziale di bias algoritmici. Analizza le specificità del Regolamento UE sull'IA, il suo approccio basato sul rischio, la definizione di "sistema di IA" e le implicazioni della classificazione "ad alto rischio" per le applicazioni in materia di SSL. Esplora i percorsi che i sistemi di IA per la SSL potrebbero seguire per evitare la classificazione ad alto rischio e sostiene un approccio interpretativo che promuova l'innovazione salvaguardando i diritti fondamentali. Presentato a Bruxelles nell'ambito delle due giornate promosse da EU-OSHA per lo scambio di buone prassi nel giorni 20 e 21 maggio 2025, il paper su cui si basa questa scheda è stato scelto come key speech per la prima sessione plenaria.

L'integrazione dell'Intelligenza Artificiale (IA) nel panorama della Salute e Sicurezza sul Lavoro (SSL) si configura come un'opportunità trasformativa, spostando le strategie per la salvaguardia del benessere dei lavoratori da reattive a proattive e predittive. L'IA, con la sua capacità di analizzare grandi volumi di dati, rilevare schemi complessi e operare con vari gradi di autonomia, offre strumenti innovativi: dall'analisi predittiva per anticipare pericoli al monitoraggio in tempo reale per rilevare rischi emergenti o stati fisiologici dei lavoratori, fino alla guida di sistemi automatizzati di prevenzione e alla personalizzazione della formazione. L'obiettivo è una significativa riduzione di incidenti, lesioni e malattie sul lavoro, orientandosi verso ambienti più intelligenti e sicuri.
Tuttavia, questo potenziale si accompagna a complessità normative ed etiche. Preoccupazioni legate alla privacy dei dati, al potenziale di bias algoritmici e all'opacità di alcuni modelli avanzati di IA ("scatola nera") rendono indispensabile un approccio normativo ponderato. La sfida è trovare un delicato equilibrio: promuovere l'innovazione senza soffocare lo sviluppo di soluzioni salvavita, ma al contempo prevenire usi impropri che potrebbero erodere la fiducia dei lavoratori e i diritti fondamentali, nonché comportare gravi sanzioni. Si tratta in pratica di gestire un carico regolamentare che potrebbe frenare l'adozione di sistemi IA in grado di migliorare la sicurezza, favorendo un approccio graduale e sostenibile ai necessari adeguamenti normativi.
Il Regolamento (UE) 2024/1689), primo quadro giuridico globale per l'IA, mira a promuovere

una tecnologia affidabile e centrata sull'uomo, tutelando salute, sicurezza e diritti fondamentali. Adotta un approccio basato sul rischio, classificando i sistemi di IA in livelli: rischio inaccettabile (proibiti, come il social scoring), alto rischio (ammessi ma soggetti a requisiti rigorosi), rischio limitato (soggetti a obblighi di trasparenza, come i chatbot) e rischio minimo (libero utilizzo). Il campo di applicazione del Regolamento è vasto, estendendosi anche a fornitori e utilizzatori al di fuori dell'UE se i loro sistemi influenzano persone o sono immessi sul mercato europeo.

La definizione di "sistema di IA" (Articolo 3(1)) è cruciale: "un sistema basato su macchina... che *inferisce*... come generare output... che possono influenzare ambienti fisici o virtuali". La Commissione Europea ha fornito linee guida non vincolanti, individuando sette elementi chiave cumulativi. Tra questi, l'operare con "vari livelli di autonomia", la capacità di "esibire adattabilità dopo l'implementazione" (sebbene non obbligatoria per la classificazione IA), la presenza di "obiettivi espliciti o impliciti", la capacità di "*inferire*... come generare output", e la possibilità di "*influenzare ambienti fisici o virtuali*". L'ampiezza di tale definizione potrebbe includere sistemi automatizzati complessi tradizionalmente non etichettati come "IA", potenzialmente introducendo obblighi normativi e costi imprevisti anche per tecnologie SSL consolidate. La distinzione tra gli "obiettivi" interni di un sistema e il suo "scopo previsto" è inoltre fondamentale per una corretta valutazione dei rischi secondo le previsioni della norma.

Una specifica ambiguità terminologica presente nel Regolamento emerge poi dalla traduzione del verbo inglese "infers" (inferisce), presente nel testo inglese e reso nelle versioni italiana e francese del Regolamento con "deduce" e in quella tedesca con "ableitet" - più corrispondente all'italiano "deriva". In pratica, mentre dobbiamo applicare in Europa il Regolamento come base di un diritto armonizzato sull'IA, in realtà oggi non possiamo sapere con certezza quale modalità operativa caratterizzi l'ambito oggettivo dei sistemi regolati dalla nuova norma. "Inferire" è un termine che si riferisce a ragionamenti induttivi e probabilistici tipici dell'apprendimento automatico. "Dedurre" implica un ragionamento più diretto e logicamente certo. "Derivare" esprime un significato ancora diverso che potrebbe sottolineare una relazione più diretta con i dati. Le differenze linguistiche potrebbero in ogni caso portare a interpretazioni divergenti e causare incertezza giuridica, influenzando la decisione su quali sistemi IA per la SSL ricadano nell'ambito del Regolamento. Per gli sviluppatori e gli utilizzatori, questa incoerenza può generare problemi di conformità transfrontaliera.

Un altro punto ambiguo riguarda le espressioni "ambiente virtuale o fisico" e "può influenzare". Le linee guida della CE chiariscono che gli ambienti fisici in questione riguardano oggetti tangibili e spazi di lavoro, mentre quelli virtuali comprendono spazi digitali e flussi di dati. L'espressione "può influenzare" suggerisce che non è necessaria una modifica sostanziale o un legame causale diretto; il *potenziale* di influenza è sufficiente. Questo potrebbe coinvolgere anche sistemi IA informativi o consultivi in SSL, il cui impatto è mediato dal giudizio umano. L'ambiguità su questa soglia di rilevanza normativa potrebbe sottoporre strumenti IA "soft" a un controllo sproporzionato.

Per i professionisti della sicurezza sul lavoro, è necessario comprendere che i sistemi di IA per la SSL sono classificati in generale come "ad alto rischio", una prospettiva prevalente nel Regolamento che vede l'IA come nuovo rischio introdotto dalla tecnologia, non come uno strumento per ridurre o eliminare rischi già presenti. Rientrano potenzialmente nel settore

"alto rischio" le applicazioni relative all'occupazione, alla gestione dei lavoratori o alle infrastrutture critiche. La classificazione ad alto rischio comporta oneri significativi: sistemi di gestione del rischio, requisiti di qualità dei dati, ampia documentazione, trasparenza, supervisione umana, accuratezza, robustezza, cybersecurity, valutazione di conformità (talvolta con organismi terzi), registrazione e monitoraggio post-commercializzazione. Questi obblighi possono tradursi in costi elevati e fungere da barriera all'innovazione, soprattutto per le PMI.
È quindi opportuno esplorare i percorsi per evitare la classificazione di un sistema destinato alla SSL come ad alto rischio. L'articolo 6, paragrafo 3, del Regolamento sull'IA prevede a questo proposito una deroga se il sistema non presenta un "rischio significativo di danno" e non "influenza materialmente l'esito del processo decisionale", a condizione che svolga un compito procedurale ristretto, migliori un'attività umana già completata, rilevi modelli decisionali senza sostituire la valutazione umana, o svolga un compito preparatorio a una valutazione. Tuttavia, un sistema di IA dell'Allegato III sarà sempre ad alto rischio se effettua profilazione di persone fisiche. L'onere della prova per la deroga ricade sul fornitore. L'ambiguità di termini come "rischio significativo di danno" e "influenzare materialmente" crea ulteriore incertezza, potendo portare a un'eccessiva cautela nella classificazione.
Le linee guida della Commissione Europea, attese entro febbraio 2026, forniranno esempi pratici. Per i sistemi IA per la SSL, la classificazione non ad alto rischio potrebbe ad esempio applicarsi a:

- **IA per l'istruzione e la formazione generale** (se non valuta i partecipanti con conseguenze significative per la loro carriera);
- **Sistemi informativi a supporto dei decisori umani**, purché l'IA non determini la decisione finale e venga preservata l'autonomia umana. È necessario a questo proposito evitare il fallimento del requisito "Human In The Loop" (HITL), che si determina se l'umano non è adeguatamente equipaggiato per valutare criticamente o ignorare l'output dell'IA
- **Sistemi IA volti puramente alla riduzione del rischio**, che non introducano nuovi rischi. La designazione di "componente di sicurezza" per prodotti soggetti ad altra legislazione UE (es. Direttiva Macchine) potrebbe tuttavia far presumere la classificazione ad alto rischio. In tal caso, il soggetto interessato all'esenzione dovrebbe dimostrare che gli aspetti specifici del sistema nel cui ambito opera l'IA restano estranei al perimetro che il regime normativo "alto rischio" intende mitigare.

Per promuovere un'IA etica ed efficace in SSL, in questa fase di avvio del quadro normativo, è essenziale un approccio interpretativo attento, basato su trasparenza, responsabilità e inclusività. La trasparenza riguarda sia il processo normativo sia l'informazione ai lavoratori sull'uso dei sistemi IA. La responsabilità implica la definizione di chi risponde in caso di errori o danni, con meccanismi di risarcimento allineati alle coperture assicurative. L'inclusività richiede il coinvolgimento di tutte le parti interessate (sviluppatori, professionisti SSL, datori di lavoro, lavoratori e loro rappresentanti), come già previsto dalla Direttiva Quadro UE sulla SSL.

Un dialogo continuo, lo sviluppo di pratiche comuni e linee guida settoriali, e la promozione della standardizzazione sono essenziali per interpretare le ambiguità del Regolamento e facilitare l'adozione sicura dell'IA.
Organismi come EU-OSHA e i soggetti nazionali competenti per la SSL giocano un ruolo vitale in questo processo. Le linee guida della Commissione Europea sull'articolo 6, paragrafo 3, saranno un passo importante. Gli sforzi devono essere sostanziali, non meramente procedurali e documentali, per evitare un "teatro della conformità".
La rapida evoluzione dell'IA impone un approccio di governance agile e adattivo, con meccanismi di apprendimento continuo e aggiornamenti normativi. Anche il lavoro degli organismi di standardizzazione è fondamentale per tradurre i principi astratti in standard tecnici verificabili.
In definitiva, la fiducia nei sistemi di IA per la SSL non è automatica, ma va guadagnata attraverso progettazione e funzionamento trasparenti, validazione robusta, supervisione umana significativa, quadri di responsabilità chiari e governance inclusiva. Il mancato raggiungimento di questo equilibrio comporta significative implicazioni economiche a lungo termine, tra cui il rischio di rimanere indietro nell'innovazione e costi sociali ed economici derivanti da incidenti o perdita di fiducia.
Le parti interessate devono agire in modo concertato:

- i regolatori fornendo linee guida chiare e adattive;
- gli sviluppatori e i fornitori adottando principi di "sicurezza ed etica fin dalla progettazione";
- i datori di lavoro conducendo valutazioni dei rischi e consultando i lavoratori, investendo nell'alfabetizzazione sull'IA e implementando una supervisione umana robusta;
- i professionisti della SSL sviluppando competenze e partecipando alla definizione delle migliori pratiche;
- i lavoratori e i loro rappresentanti partecipando attivamente e sostenendo i loro diritti;
- i ricercatori contribuendo all'evoluzione dei quadri etici e di governance.

Un approccio equilibrato e centrato sull'uomo è la chiave per sbloccare il pieno potenziale dell'IA, creando luoghi di lavoro più sicuri e salubri, e garantendo un progresso sostenibile nell'era dell'IA.