



FOCUS LAVORO, PERSONA, TECNOLOGIA
3 APRILE 2024

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

di Michele Faioli

Professore associato di Diritto del lavoro
Università Cattolica del Sacro Cuore



*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**

di Michele Faioli

Professore associato di Diritto del lavoro
Università Cattolica del Sacro Cuore

Abstract [It]: Il saggio muove da una constatazione della realtà: la trasformazione tecnologica, dovuta all'intelligenza artificiale di nuova generazione (qui anche "Frontier AI", secondo la definizione dell'AI Safety Summit di Londra del 2023), ci costringe a fare una nuova mappatura degli effettivi rischi e delle possibilità di innovazione che derivano dall'interazione tra lavoratori e macchina intelligente. Il che ci pone di fronte a domande più complesse di quelle che ci siamo posti sino a oggi. Qui, a tal fine, si sceglie l'angolo visuale dei nuovi rischi sociali e psico-fisici derivanti da tale interazione persona/macchina intelligente. Pur essendo convinti che da quella interazione possano derivare certamente maggiore produttività e maggiore benessere, non si può non indagare ciò che sta per accadere nei luoghi di lavoro ri-plasmati dalla tecnologia avanzata (Frontier AI), nei quali la nozione classica di rischio probabilmente non riesce più a essere adeguata in ragione della presenza operativa di un terzo elemento tra datore di lavoro e lavoratore. Tale terzo elemento esercita poteri, si confronta con obblighi nonché può determinare danni, creando forme di responsabilità contrattuale e extracontrattuale. Il saggio, nell'ambito di un progetto di ricerca più ampio e transdisciplinare, intende avviare il confronto accademico per poter creare un substrato teorico di una nuova branca della disciplina giuslavoristica che attiene allo studio della regolazione dell'intelligenza artificiale/robot in interazione attiva/biunivoca con lavoratori in contesti produttivi tecnologici avanzati (qui definita anche *Robot Labor Law* – "RLL").

Title: *Robot Labor Law*. Research lines for a new branch of labour law and in view of the 2024 G7 session on artificial intelligence

Abstract [En]: The tech transformation is forcing us to rethink new solutions for mapping risks and opportunities arising from interactions between robot and worker (IWRs). While doing so, we must keep in mind that the debate is not about robots and artificial intelligence, but about us, who will have to live and work with them. In such line, we are aware that there are challenging changes. Such changes should be investigated in relation to the occurrence of new social and psychophysical risks connected to IWRs. My essay is built on the conviction that robots, empowered by AI (below also "Frontier AI", notion that is line with the 2023 London AI Safety Summit), should not merely augment humans in their capabilities, but that robots and workers can become better together, thereby boosting worker wellbeing and productivity. The absence of a unified scientific framework makes us blind to various forms of reciprocity that exist when robots share the same physical and social space as workers. My thesis is aimed at buttressing the forefront of understanding and shaping the future of work regulation around robots. Also upon the results of an ongoing transdisciplinary research project, I intend to develop a basis for a new academic knowledge concerning the IWRs regulation (i.e. the "Robot Labor Law" - RLL). The RLL may bring unprecedented benefits to individual workers, teams, corporations, institutions, and organizations.

* Articolo sottoposto a referaggio. Il presente studio è stato elaborato nell'ambito delle prime attività del costituendo centro di ricerca *WSP Co-Lab, Workplace and Social Policies Co-Lab* (Università Cattolica del Sacro Cuore), del percorso di confronto con le parti sociali presso la SERI-FGB nonché in relazione alle indagini *in fieri* che l'unità di ricerca, costituita presso l'Università Cattolica del Sacro Cuore, di cui l'autore è responsabile scientifico, ha avviato con riferimento al PRIN 2022, Progetto "*SafetyChain: Social Blockchain for Implementation of a Digital Wallet for Occupational Health and Safety Training*", codice progetto "20225T2B2P", finanziato dall'Unione europea – *Next Generation EU*.

Parole chiave: relazioni di lavoro, rischio, danni, responsabilità, robot, intelligenza artificiale

Keywords: work relations, risks, damages, liability, robot, AI

Sommario: **1.** Responsabilità, rischio da lavoro alla persona, imputabilità alla macchina intelligente di nuova generazione (*Frontier AI*). Il quadro tecnologico in cui la ricerca si colloca (*Foundation Models, LLMs, Frontier AI, Generative AI*). **2.** Come si può comparare il diritto statunitense con quello euro-unitario che regola lavoro e intelligenza artificiale. Genotipi e fenotipi legali. **3.** Rischio da lavoro, danno alla persona e norma europea. Le domande da porsi. **4.** Rischio, danno alla persona e norma statunitense. Cenni. **5.** Conclusioni. Per impostare i contenuti di una nuova branca della disciplina giuslavoristica (RLL - *Robot Labor Law*).

Direction is more important than speed.

R. Feynman

*The Vantaser Danger Meridian is the
point or line after which danger to
your mission and/or sense of self increases exponentially.*

*Often used to demarcate
conditions of grave and approaching danger...*

S. Conrad

1. Responsabilità, rischio da lavoro alla persona, imputabilità alla macchina intelligente di nuova generazione. Il quadro tecnologico in cui la ricerca si colloca (*Foundation Models, LLMs, Frontier AI, Generative AI*)

Con il presente studio si intende avviare un dialogo, anche con i colleghi dell'accademia, giuristi e esperti di tecnologia, nonché con sindacati e organizzazioni datoriali, sull'adeguatezza degli attuali schemi giuridici che permettono di realizzare prevenzione, mitigazione e protezione dei nuovi rischi da lavoro riferibili all'intelligenza artificiale e alla robotica, operanti a livello di unità produttiva, anche in vista delle scelte di politica del diritto che potranno essere effettuate in materia a seguito del G7 del 2024.

Le domande da cui muoviamo sono le seguenti: cosa desideriamo che l'intelligenza artificiale, anche nella forma di robotica avanzata, nei luoghi di lavoro, faccia per noi, o, meglio, "con" noi, e certamente mai a nostro danno? Chi regola questa interazione tra la persona del lavoratore e l'intelligenza artificiale/*robot* (qui anche "AI/R")? E come si deve regolare quell'interazione? Inoltre, data tale interazione, cosa potrebbe accadere qualora la responsabilità dell'eventuale danno alla persona del lavoratore dipendesse direttamente/esclusivamente dall'AI/R? L'assicurazione obbligatoria contro gli infortuni sul lavoro e le tecnopatie presenta già oggi una fisionomia per poter comprendere nel nostro sistema previdenziale anche quel danno da lavoro dipendente direttamente da AI/R? E quanto è efficiente, in termini di mitigazione del danno, la procedimentalizzazione che l'attuale regolazione sulla sicurezza sul lavoro impone, anche nell'ambito di produzioni a tecnologia avanzata o dove il lavoratore co-opera con l'AI/R?

Tali domande, come si intuisce, non attengono alla mancanza o alla insufficienza della tecnologia, ma a problemi che sono nostri perché totalmente umani, riferibili alle regole che ci diamo, sia di livello individuale che collettivo. Per poter incidere in positivo sui problemi che originano da tali domande, riferite alla regolazione della tecnologia avanzata, AI/R, nei luoghi di lavoro, si deve preliminarmente osservare il campo dell'umano, chiedersi cosa intendiamo fare noi per migliorare quel campo e, infine, quale obiettivo si intende davvero raggiungere con la regolazione e la verifica della conformità a quella regolazione.

Le osservazioni che seguono vanno ovviamente oltre l'ambito di analisi etica, il quale è pressoché orientato, anche a livello internazionale, a discernere tra ciò che è super-potere dell'AI/R e ciò che non lo è, e, di conseguenza, tra ciò che è bene e ciò che è male per l'umanità nella creazione di questo super-potere dell'AI/R¹. Qui, si vorrebbe, invece, disaminare la prospettiva della regolazione di ciò che non si conosce ancora pienamente e, in particolare, della regolazione dell'AI/R nella cooperazione attiva con il lavoratore in ambienti altamente tecnologici, composti da nuovi modelli di intelligenza artificiale (*Frontier AI*), tenendo in considerazione che il quadro generale socio-economico e industriale, in cui questo contributo scientifico si pone, può forse essere comparato con quello dell'avvio delle sperimentazioni dell'energia nucleare e, più recentemente, della grande crisi finanziaria del 2008². In entrambi quei casi la regolazione macro-regionale, nazionale, domestica, è servita a poco o niente. Forse è arrivata anche troppo tardi e si è mostrata inefficiente. Si è dovuto, invece, trasferire a un regolatore internazionale la capacità di creare un governo complessivo dell'intera vicenda (durante e dopo la Seconda guerra mondiale per il nucleare e dal 2008 in poi per la crisi finanziaria). I risultati, in quei casi, sono stati parziali, non definitivi, forse non soddisfacenti, ma certamente orientati verso la direzione giusta: regolare a livello nazionale il fenomeno in sé serve a poco, mentre può essere più utile regolare le condotte di chi quel fenomeno deve gestire/governare a livello nazionale, imponendo *standards* internazionali e creando relative autorità di vigilanza. Il che, in qualche modo, sta già accadendo attorno alla definizione di AI/R che l'OCSE ha recentemente individuato, anche per avviare delle forme di cooperazione tra diverse giurisdizioni³. In tale prospettiva, i sistemi giuridici occidentali, a livello transatlantico⁴, si stanno

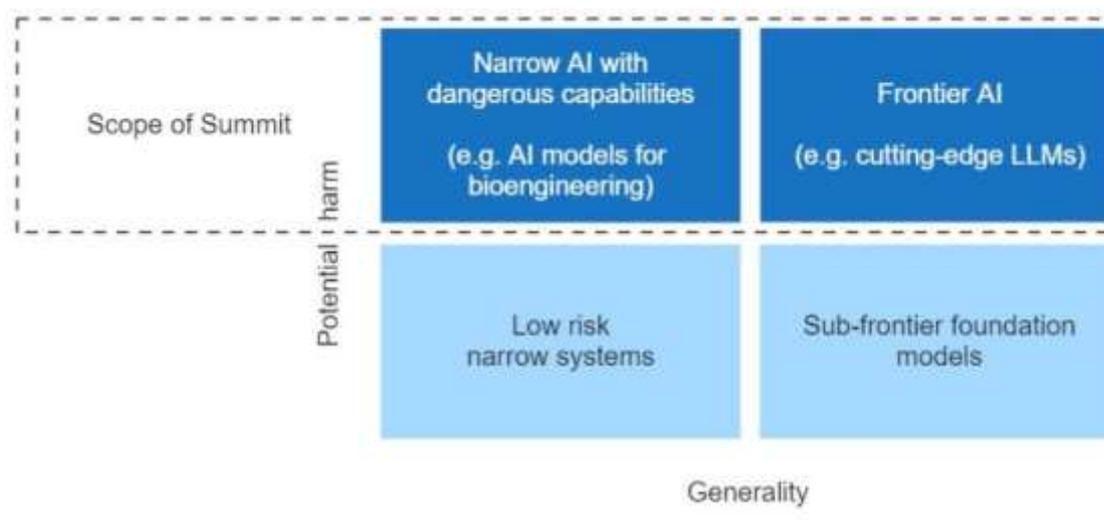
¹ Si v. i lavori della Commissione ONU sugli aspetti etici dell'intelligenza artificiale, *AI Advisory Board* - <https://www.un.org/en/ai-advisory-body>. Per le indagini accademiche recenti più importanti si rinvia al libro di D. J. GUNKEL, *Person, Thing, Robot. A Moral and Legal Ontology For the 21st Century And Beyond*, MIT Press, Cambridge, 2023, il quale richiama anche gli studi fondativi di P. RICOEUR, *Il giusto*, Effatà, Torino, 1-2, 2005 e di R. ESPOSITO, *Le persone e le cose*, Einaudi, Torino, 2014.

² L'intuizione è di L. ZINGALES, B. MCLEAN, *Who Controls AI? With Sendhil Mullainathan*, in *Capitalism't*, 21 dicembre 2023.

³ Si v. la documentazione qui riportata - <https://oecd.ai/en/wonk/ai-system-definition-update>

⁴ Gli esiti del dialogo tra Unione europea e Stati Uniti d'America si possono analizzare in https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/stronger-europe-world/eu-us-trade-and-technology-council_en#objectives-of-the-partnership

orientando anche per l'intelligenza artificiale. In particolare, il diritto euro-unitario e quello statunitense stanno cercando risposte, anche per regolare la complessità del fenomeno, in base a un presupposto giuridico che si evolve assai velocemente e deve valere a livello transnazionale: si cerca di porre una nozione giuridica di AI/R che si mostri più adatta possibile al contesto futuro, non a quello che già sappiamo oggi o è accaduto nel passato. Tale nozione coincide, almeno in questo frangente storico, in Europa e negli Stati Uniti d'America, con quella di “*Frontier AI Models*”, qui anche intelligenza artificiale di nuova generazione o di frontiera⁵. Dal punto di vista tecnologico possiamo definire *Frontier AI* ciò che rientra nella struttura dei cd. *foundation models*, i quali normalmente possiedono altresì una certa capacità di determinare danni. Rientrano nella definizione di *foundation model* anche *LLMs*, *BERT*, *DALL-E*, *GPT-3*. Si tratta di un modello che è costruito su una larga base di dati ed è adattabile a una serie quasi infinita di compiti (cd. *downstream tasks*). Lo schema che segue è tratto dai documenti dell'*AI Safety Summit* di Londra, 2023.



In relazione a tale definizione, a noi interessa mettere al centro delle riflessioni giuslavoristiche sull'AI/R che agisce in cooperazione con il lavoratore il modo mediante cui, da una parte, si assicura, secondo il diritto della sicurezza sociale, e, dall'altra, si può mitigare, con strumenti di prevenzione da sicurezza sul lavoro, quel danno che può derivare anche da *downstream task* effettuato nei luoghi di lavoro, nell'ambito di interazioni tra persona e macchina, gestita da una *Frontier AI*.

⁵ Per la definizione di *Frontier AI* si v. i risultati della conferenza intergovernativa “*AI Safety Summit*”, Londra, 2023, e, in particolare, il documento *Frontier AI: Capabilities and Risks – Discussion Paper*, 25 ottobre 2023 in <https://www.gov.uk/government/publications/frontier-ai-capabilities-and-risks-discussion-paper> - “*For the purposes of the Summit we define frontier AI as highly capable general-purpose AI models that can perform a wide variety of tasks and match or exceed the capabilities present in today’s most advanced models. Today, this primarily includes large language models (LLMs) such as those underlying ChatGPT, Claude, and Bard. However, it is important to note that, both today and in the future, frontier AI systems may not be underpinned by LLMs, and could be underpinned by another technology*”.

Il saggio muove, pertanto, da una constatazione della realtà: la trasformazione tecnologica, dovuta all'intelligenza artificiale di nuova generazione (*Frontier AI*), ci costringe a fare una nuova mappatura degli effettivi rischi e delle possibilità di innovazione che derivano dall'interazione tra lavoratori e macchina intelligente. Il che ci pone di fronte a domande più complesse di quelle che ci siamo posti sino a oggi. A tal fine si sceglie l'angolo visuale dei nuovi rischi sociali e psico-fisici derivanti da tale interazione persona/macchina intelligente. Pur essendo convinti che da quella interazione possano derivare certamente maggiore produttività e maggiore benessere, non si può non indagare ciò che sta per accadere nei luoghi di lavoro ri-plasmati dalla tecnologia avanzata (*Frontier AI*), nei quali la nozione classica di rischio probabilmente non riesce più a essere adeguata in ragione della presenza operativa di un "terzo elemento" tra datore di lavoro e lavoratore⁶. Tale terzo elemento esercita poteri, si confronta con obblighi nonché può determinare danni, creando anche forme di responsabilità contrattuale e extracontrattuale. Il saggio, nell'ambito di un progetto di ricerca più ampio e transdisciplinare, intende avviare il confronto, anche accademico, per poter creare un substrato teorico di una nuova branca della disciplina giuslavoristica che attiene allo studio della regolazione dell'intelligenza artificiale/*robot* in interazione attiva/biunivoca con lavoratori in contesti produttivi tecnologici (qui definita anche *Robot Labor Law* – "RLL"). Il saggio, definendo il metodo comparativo prescelto (paragrafo 2), imposta l'analisi dell'assetto delle regole europee e del relativo contesto giuslavoristico, i quali, con buona probabilità, avranno come primo compito quello di definire il tipo di rischio derivante dalla *Frontier AI*, anche nella relativa operatività nei luoghi di lavoro (paragrafo 3). Si continua con l'analisi dell'attuale sistema statunitense di regolazione della *Frontier AI*, il quale pare abbia come scopo anche il trascinarsi geo-politico su scala globale della normazione della *Frontier AI* (paragrafo 4). Nel paragrafo 5 si proverà a delineare qualche contenuto della nuova branca, di rilievo transnazionale, della disciplina giuslavoristica "*Robot Labor Law*" – RLL, in attesa del completamento del progetto di ricerca che sottende questo saggio.

2. Come si può comparare il diritto statunitense con quello euro-unitario che regola lavoro e intelligenza artificiale. Genotipi e fenotipi legali

Il metodo per la comparazione tra diritto euro-unitario e diritto statunitense può muovere dalla costruzione di un genotipo, da cui, in un secondo momento, far discendere i fenotipi giuridici specifici,

⁶ Sulla nozione di macchina intelligente come "terzo elemento" del rapporto di lavoro rinvio ai miei precedenti studi. In particolare, si v. M. FAIOLI, *Mansioni e macchina intelligente*, Torino, 2018 e la linea teorica definita con alcuni miei scritti, tra cui in particolare M. FAIOLI, *Data analytics, robot intelligenti e regolazione del lavoro*, in q. *Rivista*, n. 9, 2022, pp. 149 ss.; M. FAIOLI M., *Artificial Intelligence: The Third Element of the Labour Relations*, in A. PERULLI, T. TREU, (a cura di), *The Future of Work. Labour Law and Labour Market Regulation in the Digital Era*, Alphen aan den Rijn, Wolters-Kluwer, 2021; M. FAIOLI, *Unità produttiva digitale. Perché riformare lo Statuto dei lavoratori*, in *Economia & lavoro, Rivista di politica sindacale, sociologia e relazioni industriali*, n. 1, 2021, pp. 41 ss.

validi per osservare i due sistemi legali⁷. Il genotipo si può costruire in relazione a tre problemi: (i) quale rischio e quali danni quell'ordinamento giuridico intende selezionare ai fini della tutela da predisporre in relazione alla capacità della *Frontier AI* di interagire con il lavoratore? (ii) Quali misure di mitigazione e prevenzione quell'ordinamento giuridico può introdurre per far fronte al bisogno di tutela correlato a quei rischi e a quei danni? (iii) Quali istituzioni quell'ordinamento giuridico intende porre per attuare tale tutela, anche in termini di esigibilità individuale e collettiva?

I tre problemi ci permettono di comprendere in quale misura il legislatore sia più orientato verso una norma di prevenzione del rischio o, al contrario, verso una norma punitiva, cioè più orientata a punire il danno e a fissare la relativa riparazione. Sappiamo che la nozione di rischio, individuale e collettivo, non è mai neutra, almeno se osservata con gli occhi di un giurista. Essa è frutto di una scelta di politica del diritto ben precisa, in questo caso effettuata a livello europeo e statunitense. Aver deciso di collocare nella nozione di rischio quei (potenziali) danni alla persona derivanti da condotta di AI/R, anche nella specie di *Frontier AI*, significa aver predeterminato un percorso di regolazione, il quale non coincide più (o non coincide solo) con la richiesta di risarcimento di un danno da parte della persona che lo subisce, ma soprattutto con una norma che utilizza un certo modo di regolare la gestione del rischio, in continuità o analogamente a altri settori ove esistono prassi già consolidate di gestione del rischio (energia, ambiente, etc.). Anche nel sistema dell'intelligenza artificiale, si può decidere di regolare *ex ante* tali danni, facendo ricorso alla nozione di rischio, con modelli anticipatori, di mitigazione o precauzionali, introducendo una specie di trapianto di istituti giuridici già utilizzati altrove e probabilmente utili anche in questo caso.

Di qui si può ritenere che il diritto dimostri di partecipare attivamente, quasi in una logica di costruzione concettuale, alla definizione di *Frontier AI*, proprio in ragione del tipo di rischio (individuale e collettivo) che si intende selezionare e, poi, in ragione della norma, mitigare, prevenire, assicurare, etc.

⁷ Qui si segue il metodo elaborato, presso la *Cornell Law School*, da R. B. SCHLESINGER, *Comparative Law. Cases, Text, Materials*, Foundation Press, New York, edizione del 1988, da G. GORLA, *Diritto comparato e diritto comune europeo*, Giuffrè, Milano, 1981 e da R. SACCO, *Introduzione al diritto comparato*, Giappichelli, Torino, 1990. C'è un riferimento agli studi di L. MENGONI, *Diritto vivente*, in *Jus*, n. 1, 1988, pp. 14 ss. e di G. BENEDETTI, *L'elogio dell'interpretazione traducevole nell'orizzonte del diritto europeo*, in *Europa e diritto privato*, n. 2, 2010, pp. 413 ss.. Si v. anche la raccolta di saggi in P. SANDULLI, M. FAIOLI (a cura di), *Attività transnazionali. Sapere giuridico e scienza della traduzione*, Edizioni Nuova Cultura, Roma, 2011.

Il rischio segna una direzione, un orientamento, e, per alcuni versi, un dato da osservare nel futuro⁸. Il danno è, invece, già accaduto⁹. Il diritto reagisce rispetto al rischio con una serie di manovre preventive. In relazione al danno, il diritto pone, invece, tutele riparatorie in ragione di un fatto che ha inciso negativamente sulla persona, fissando le regole sul “se”, sul “chi” e sul “quanto/come” (se il danno verificatosi debba essere risarcito, chi è obbligato a risarcirlo, quanto/come si debba riparare il danno)¹⁰. Dove si regola il rischio, sussumendo in un certo senso il danno all’interno della nozione di rischio, si compie una mera operazione descrittiva, basata su statistiche, probabilità, campioni. In ragione di ciò si indicano i modi mediante cui si può ridurre la possibilità del verificarsi di quel rischio, imponendo auto-verifiche, *due diligence*, richieste di autorizzazione/permessi, etc.

In generale, il rischio ha in sé un certo meccanismo di selezione dell’evento da prevenire, assicurare, mitigare, etc. perché la regola stessa sul rischio si basa su un calcolo per prevenire, assicurare, mitigare il rischio medesimo. Nel caso dell’intelligenza artificiale, ciò è tanto più vero perché è la stessa legge che fissa o aggiorna la nozione di intelligenza artificiale, muovendo da una variabilità indefinita di definizioni tecnologiche, individuandone una o alcune per gli effetti giuridici desiderati in quel momento storico e in quel contesto socio-politico. Ed è in relazione a tale nozione di intelligenza artificiale (per esemplificare, oggi è la *Frontier AI*, domani sarà altro) che scattano una serie di meccanismi di protezione e di obblighi di comportamento.

Il rischio potenziale derivante dall’AI/R, anche gestita dalla *Frontier AI*, è quasi sempre contestuale. Esso dipende dal contesto operativo in cui l’AI/R è posta. La mappatura del rischio non può essere effettuata una volta per sempre perché il rischio muta in ragione di una serie di fattori che interagiscono tra essi,

⁸ La nozione di rischio è stata oggetto di importanti studi giuslavoristici e di diritto della sicurezza sociale. Ci sono gli studi confluiti nella collettanea AIDLASS, *Rischio e bisogno nella crisi della previdenza sociale*, Giuffrè, Milano, 1985, la quale permette di approfondire alcuni dei temi trattati da F. SANTORO PASSARELLI, *Rischio e bisogno nella previdenza sociale*, Giuffrè, Milano, 1948. Tra le analisi più recenti c’è quella svolta da P. LOI, *Il principio di ragionevolezza e proporzionalità nel diritto del lavoro*, Giappichelli, Torino, 2017. Della medesima A. si v. anche *Il rischio proporzionato nella proposta di regolamento sull’IA e i suoi effetti nel rapporto di lavoro*, in *q. Rivista*, n. 4, 2023, pp. 239 ss. Si v. anche gli spunti che si trovano nei saggi di T. TREU, *Il diritto del lavoro: realtà e possibilità*, in *ADL Argomenti di diritto del lavoro*, n. 3, 2000, pp. 467 ss.; S. PICCININNO, P. SANDULLI, *Le riforme della sicurezza sociale nell’Europa occidentale*, in *Il Diritto del lavoro*, n. 3-4, 1980, pp. 199 ss.; M. DELL’OLIO, *Previdenza privata e libertà*, in *ADL Argomenti di diritto del lavoro*, n. 2, 1998, pag. 491 ss.

⁹ Si v. la rilettura nell’ambito del diritto civile di alcuni elementi teorici di base della nozione di danno effettuata recentemente in G. ALPA, *Danno "in re ipsa" e tutela dei diritti fondamentali (diritti della personalità e diritto di proprietà)*, in *Responsabilità civile e previdenza*, n. 1, 2023, pp. 6 ss.; P. SIRENA, *Danno-evento, danno-conseguenza e relativi nessi causali. Una storia di superfetazioni interpretative e ipocrisie giurisprudenziali*, in *Responsabilità civile e previdenza*, n. 1, 2023, pp. 68 ss.; V. ROPPO, *Pensieri sparsi sulla responsabilità civile (in margine al libro di Pietro Trimarchi)*, in *Questione Giustizia*, n. 1, 2018, pp. 108 ss.; P. GALLO, *Quale futuro per il contatto sociale in Italia?*, in *La Nuova Giurisprudenza Civile Commentata*, n. 12, 2017, pp. 1759 ss. V. ROPPO, *Responsabilità contrattuale: funzioni di deterrenza?*, in *Lavoro e diritto*, n. 3-4, 2017, pp. 407 ss. Si v. anche i recenti studi giuslavoristici in materia di danno di A. BOLLANI, *Il danno alla persona nel diritto del lavoro, tra influssi della civilistica e necessari adattamenti*, in *Giornale di diritto del lavoro e di relazioni industriali*, n. 176, 2022, pp. 593 ss.

¹⁰ Per una panoramica sul dibattito civilistico in corso si v. i recenti studi di A. D’ADDA, *Danni "da robot" (specie in ambito sanitario) e pluralità di responsabili tra sistema della responsabilità civile ed iniziative di diritto europeo*, in *Rivista di diritto civile*, n. 5, 2022, pp. 805 ss. Si v. anche G. CALABRESI, E. AL MUREDEN, *Driverless cars. Intelligenza artificiale e futuro della mobilità*, Il Mulino, Bologna, 2021.

con la conseguenza che è rischio ciò che in quel momento può essere definito tale. Il quadro dei fattori può cambiare e, di conseguenza, cambia anche il rischio, il quale, forse, in alcuni casi, può subire una specie di declassamento in altro. Inoltre, c'è un rischio derivante dall'AI/R, gestita dalla *Frontier AI*, solo se con il relativo utilizzo si incrementa l'esposizione al danno: il che, almeno nella logica giuslavoristica di questo studio, significa osservare il rischio con riferimento alle persone che interagiscono anche nei luoghi di lavoro con forme di AI/R, gestita dalla *Frontier AI*. In tale prospettiva, più forte sarà il contesto umano, perché emancipato, educato, formato a gestire i problemi eventuali di quell'AI/R, minore sarà l'impatto del rischio.

In tale prospettiva i tre problemi posti sopra (quali rischi, quali misure di mitigazione, quali istituzioni) diventano una specie di schema generale per poter più efficacemente comparare il diritto euro-unitario con quello statunitense, avendo a riguardo, da una parte, il diritto della tecnologia, così come si sta sviluppando in questi ultimi tempi, e, dall'altra, il diritto del lavoro che reagisce (come noto, da sempre) alla tecnologia introdotta a livello di unità produttiva in relazione a più fattori, oltre a quelli già in qualche modo oggetto di indagine (direzione, coordinamento, controllo, variabilità delle mansioni, indagini sulle opinioni, discriminazioni, etc.)¹¹.

C'è, infatti, un campo elettivo per svolgere questa disamina: quello della sicurezza sui luoghi di lavoro, dei danni alla persona e della relativa assicurazione (INAIL), intendendo quella tecnologia avanzata (*Frontier AI*) sia come strumento da cui può dipendere un danno alla persona del lavoratore, in ragione di una violazione delle norme sulla sicurezza, sia come strumento che aiuta a prevenire il danno alla persona. Con altre parole, si intende osservare, mediante il progetto di ricerca che si è avviato, la *Frontier AI* nell'ambito delle proprie funzioni di terzo elemento, il quale, in ragione del fatto che può coordinare, controllare, dirigere anche il lavoro umano, variandone le mansioni e teoricamente applicando sanzioni disciplinari, esso è, da una parte, capace di causare il danno e, dall'altra, almeno auspicabilmente, anche di prevenirlo.

Da ciò deriva la centralità dell'oggetto del presente studio: sapendo che il diritto è chiamato a adeguarsi alla tecnologia¹², non il contrario, non basta più (solo) comprendere quali limiti porre, mediante legge e/o contrattazione collettiva, rispetto ai poteri che la *Frontier AI* può esercitare¹³, ma si deve altresì cogliere quali danni e quali rischi possono derivare dalle condotte/scelte della *Frontier AI*, operate nell'interazione con la persona umana nei luoghi di lavoro, e, dunque, quali criteri di imputazione della responsabilità siamo tenuti a aggiornare e perché, dato il contesto produttivo che si evolve giorno per giorno, nonché

¹¹ Si v. la bibliografia riportata nella nota 7 e 23.

¹² Per evocare una delle idee più forti di G. GIUGNI, *Il processo tecnologico e la contrattazione collettiva*, in F. MOMIGLIANO, *Lavoratori e sindacati di fronte alle trasformazioni del processo produttivo*, Feltrinelli, Milano, 1962.

¹³ A tal proposito si v. la struttura della teoria in M. FAIOLI, *op. cit.*, 2018, pp. 93 ss. e pp. 211 ss.

come prevenire e mitigare tali rischi e danni, anche mediante innovati procedimenti organizzativi di livello aziendale e avanzati sistemi tecnologici che possono essere introdotti con la specifica funzione di prevenzione/mitigazione dei rischi/danni da *Frontier AI*.

3. Rischio da lavoro, danno alla persona e norma europea. Le domande da porsi

Normalmente, di fronte a fenomeni così complessi come quelli connessi ai danni alla persona del lavoratore imputabili a una macchina intelligente, il giurista suggerisce di intervenire con una norma. L'economista, con buona probabilità, ritiene sufficiente introdurre un prezzo per l'inefficienza tale da indurre chi deve conformarsi alla norma a farlo¹⁴. Ci sono, però, giuristi che hanno combinato le due prospettive. Tra questi ci sono certamente, con visioni diverse, G. Calabresi¹⁵ e R. Posner¹⁶, i quali hanno applicato la logica dei costi e dei benefici al diritto: si fissa il punto di partenza (nel nostro caso, i danni alla persona del lavoratore imputabili alla *Frontier AI* – “S”); poi, si stabilisce l'obiettivo dell'analisi (minimizzare i costi sociali – “C”); da ultimo, con passaggi vari, si seleziona il modo più efficiente economicamente (“V”) per raggiungere quel risultato (“R”).

Ora, utilizzando uno schema di esposizione in linea con tale teoria, potremmo avere quanto segue: per realizzare R (risultato), muovendo da S (situazione), si deve calcolare C (differenza tra costi) e identificare V (modo efficiente)¹⁷. Adattiamo il discorso al caso di cui ci occupiamo qui: C, il costo dei danni imputabili alla *Frontier AI* né può essere scaricato in alcun modo sui lavoratori, né può essere integralmente internalizzato dal datore di lavoro, trattandosi di danni derivanti non dal datore di lavoro, ma dalla *Frontier AI* che qui assumiamo essere il terzo elemento del rapporto di lavoro. Il risultato R non coincide certamente con il cercare di evitare del tutto tali danni perché ciò costerebbe troppo. E', invece, molto più sensato ritenere che si debbano ridurre il più possibile tali danni, nella misura massima secondo la tecnologia più appropriata applicabile (per l'ordinamento italiano, si v. i presupposti dell'art. 2087 c.c., anche in relazione al d.lgs. 9 aprile 2008, n. 81). Il modo efficiente (V) potrà essere realizzato, in alcuni casi, mediante una prevenzione generale e, in molti altri, mediante una prevenzione specifica, definita con provvedimenti che sono volti a identificare il modo più sicuro per far operare macchine intelligenti nei luoghi di lavoro. La prevenzione generale si basa sull'idea che il datore di lavoro possa assumersi i costi per l'introduzione dell'AI/R a livello di unità produttiva e per la relativa *compliance* rispetto a alcuni

¹⁴ R. COASE, *The Problem of Social Cost*, in *Journal of Law & Economics*, n. 3, 1969, pp. 1 ss.

¹⁵ G. CALABRESI, *Transaction Costs, Resource Allocation and Liability Rules. A Comment*, in *Journal of Law & Economics*, n. 11, 1968, pp. 67 ss.; G. CALABRESI, *Melamed, Property Rules, Liability Rules and Inalienability. A view of the Cathedral*, in *Harvard Law Review*, n. 85, 1972, pp. 1089 ss.

¹⁶ R. POSNER, *Economic Analysis of Law*, Aspen, Boston, 1992.

¹⁷ Si v. la ricostruzione teorica di E. AL MUREDEN, *Costo degli incidenti e responsabilità civile quarant'anni dopo. Attualità e nuove prospettive nell'analisi economico-giuridica di Guido Calabresi*, in *Rivista di diritto civile*, n. 4, 2015, pp. 1026 ss.

parametri fissati dalla norma. La prevenzione generale ha come presupposto l'identificazione del rischio e dei modelli di gestione di tale rischio, in relazione al livello di tollerabilità (individuale/collettivo) che si intende delineare, alla verifica che ciascuno può effettuare sul rischio che deriva da una certa condotta (in questo caso dell'AI/R) e alla relativa valutazione (rischio alto, medio, basso), nonché alle azioni di vigilanza compiute da autorità pubbliche. La prevenzione generale pone una serie di criticità quando, come nel caso dell'AI/R, c'è da fare un contro-bilanciamento con i diritti della persona umana, quelli più importanti, fondamentali, che attengono alla tutela della dignità. La prevenzione speciale viene strutturata secondo il modello di responsabilità (tendenzialmente oggettiva) di chi determina il danno, con una assicurazione che copre i costi, assicurazione che può essere affidata al mercato, o, come accade in molti paesi europei, a un gestore pubblico (per l'Italia, l'INAIL).

Tale teoria generale è adatta a interpretare e, in qualche modo altresì sottende, la norma euro-unitaria, la quale è prevalentemente impostata in vista di una prevenzione generale, cioè in ragione dei rischi da proteggere e facendo riferimento a una certa nozione variabile di AI/R (si v. proposta di regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale - legge sull'intelligenza artificiale qui di seguito "Regolamento"). L'approccio al rischio è stato regolato secondo uno schema già noto in Europa, coincidente in larga parte con i rischi derivanti dalla circolazione commerciale di prodotti o con rischi che attengono all'ambiente, alla produzione di energia, etc.. La prevenzione speciale, quella relativa ai danni da risarcire, è assegnata a una direttiva (si v. la proposta di direttiva del parlamento europeo e del consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale - direttiva sulla responsabilità da intelligenza artificiale - COM/2022/496 final – qui di seguito "Direttiva"). Il Regolamento e la Direttiva appartengono a una strategia normativa europea molto più ampia sull'intelligenza artificiale, la quale è composta da almeno quattro ambiti di regolazione (intelligenza artificiale, governo dei dati, mercati e servizi digitali, piattaforme digitali¹⁸).

Con riferimento alla prevenzione generale, il Regolamento dispone un quadro che è disegnato secondo una logica che vede, da una parte, macro-rischi riferibili al mercato e ai relativi attori (imprese/consumatori) e, dall'altra, micro-rischi che chi opera nel mercato dell'AI/R è chiamato a rilevare sia con una verifica preventiva che con analisi successive all'introduzione in quel mercato dell'AI/R. I rischi derivanti dall'AI/R sono prevalentemente parametrati sulla salute delle persone, sulla sicurezza e

¹⁸ Si v. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age_en - Si v. gli studi di M. ALMADA, N. PETT, *The EU AI Act: Between Product Safety and Fundamental Rights*, 2023, in SSRN: <https://ssrn.com/abstract=4308072> e N. MORENO BELLOSO, N. PETT, *The EU Digital Markets Act (DMA): A Competition Hand in a Regulatory Glove*, 2023, in SSRN: <https://ssrn.com/abstract=4411743>; A. BARTOLINI, *Artificial Intelligence and Civil Liability*, 2020 in [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU\(2020\)621926_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf)

sui diritti fondamentali. Ci sono rischi inaccettabili (prima tipologia)¹⁹, poi rischi elevati, ma accettabili (seconda tipologia), date alcune condizioni, e infine rischi del tutto accettabili perché minimi (terza tipologia). Ciò che può condurre ai rischi inaccettabili è assoggettato a un divieto assoluto. Ciò che determina rischi elevati viene sottoposto a un sistema di procedimenti preventivi e *ex post* di gestione. Per ciò che implica livelli minimi di rischi si promuovono varie forme di auto-regolazione. Per i rischi elevati, quelli della seconda tipologia²⁰, l'AI/R viene sottoposto *ex ante* a una verifica di conformità, ottenuta la quale si può procedere con l'immissione nel mercato. Tale verifica viene svolta in ragione del fatto che l'AI/R può essere una componente di sicurezza di un prodotto (o essa stessa un prodotto), o, in alternativa, ha riflessi sui diritti fondamentali²¹. Nel primo caso (componente di sicurezza/prodotto) c'è un ruolo di soggetti terzi che hanno poteri di notifica, ispezioni, etc. Nel secondo caso (diritti fondamentali) non c'è un ruolo di soggetti terzi che facilitano il processo di adeguamento perché c'è un divieto assoluto di ricorso a quel tipo di intelligenza artificiale. In via generale, la verifica *ex ante* muove dall'identificazione dei rischi e dalla valutazione probabilistica delle verificarsi di essi, in condizioni di normale prevedibilità. Il che viene posto anche in relazione con la valutazione di altri eventuali rischi derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione nel mercato e con l'introduzione di misure di gestione dei rischi. A valle di ciò, si chiede la conformità, anche per farsi autorizzare alla libera circolazione dell'AI/R nel mercato europeo. La conformità si basa su un atto autorizzatorio che può essere ottenuto, a seguito delle verifiche *ex ante*, anche con l'ausilio di terzi esperti. C'è altresì l'obbligo di svolgere alcune verifiche *in itinere* e *ex post* per permettere l'eventuale adozione di misure di mitigazione e la gestione continuativa nel tempo del rischio derivante dall'AI/R. Per i profili sostanziali, la norma europea impone una serie di protezioni collegate alla raccolta e all'elaborazione dei dati mediante cui si alimenta il sistema AI/R che viene immesso nel mercato. Per le verifiche *ex post* viene definito un sistema di registri pubblici, monitoraggio, segnalazione e riparazione dell'eventuale causa che

¹⁹ AI/R che utilizza tecniche subliminali, o che sfrutta le vulnerabilità di uno specifico gruppo di persone, creando discriminazione, o utilizzata da parte delle autorità pubbliche o per loro conto ai fini della valutazione o della classificazione dell'affidabilità delle persone fisiche, o capace di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico a fini di attività di contrasto ai reati, date alcune condizioni.

²⁰ I sistemi di intelligenza artificiale a alto rischio sono quelli destinati ad essere utilizzati come componenti di sicurezza di prodotti soggetti a valutazione della conformità *ex ante* da parte di terzi e quelli che presentano implicazioni principalmente in relazione ai diritti fondamentali esplicitamente elencati nell'allegato III del Regolamento. Tale elenco di AI/R ad alto rischio di cui all'allegato III del Regolamento contiene un numero di AI/R, i cui rischi si sono già concretizzati o potrebbero concretizzarsi nel prossimo futuro. La Commissione, però, può estendere l'elenco degli AI/R a alto rischio utilizzati all'interno di alcuni settori predefiniti, applicando una serie di criteri e una metodologia di valutazione dei rischi.

²¹ Tra i quali si possono annoverare per certo, dato il richiamo nel preambolo del Regolamento, il diritto al rispetto della vita privata e alla protezione dei dati di carattere personale, alla non discriminazione e alla parità tra donne e uomini, alla libertà di espressione e alla libertà di riunione, ad assicurare la tutela del diritto a un ricorso effettivo e a un giudice imparziale, della presunzione di innocenza e dei diritti della difesa, i diritti dei lavoratori a condizioni di lavoro giuste ed eque, un livello elevato di protezione dei consumatori, i diritti del minore, il diritto a un livello elevato di tutela dell'ambiente e al miglioramento della sua qualità, anche in relazione alla salute e alla sicurezza delle persone.

crea quel rischio. I sistemi AI/R che possono determinare rischi specifici di manipolazione sono sottoposti a obblighi di trasparenza specifici qualora essi interagiscano con gli esseri umani, siano utilizzati per rilevare emozioni, possano definire categorie sociali sulla base di dati biometrici, possano generare contenuti falsi.

In relazione alle istituzioni, si noti che verrà costituito un comitato europeo per l'intelligenza artificiale, con i rappresentanti degli Stati membri e della Commissione. A livello nazionale, gli Stati membri dovranno designare autorità competenti a controllare l'applicazione e l'attuazione del Regolamento. Il garante europeo della protezione dei dati agirà come autorità competente per la vigilanza delle istituzioni, delle agenzie e degli organismi dell'Unione. Ci sarà anche la creazione di una banca dati a livello europeo per i sistemi cd. ad alto rischio che presentano principalmente implicazioni in relazione ai diritti fondamentali. La banca dati sarà gestita dalla Commissione e alimentata con i dati messi a disposizione dai fornitori di AI/R, i quali saranno tenuti a registrare i relativi AI/R prima di immetterli sul mercato.

Invece, con riferimento alla prevenzione specifica (risarcimento dei danni), la Direttiva, nella consapevolezza che gli ordinamenti domestici non sono “adatti” a gestire le azioni di responsabilità per danni causati da prodotti e servizi basati sull'AI/R, perché le caratteristiche dell'AI/R, date dalla relativa complessità, l'autonomia e l'opacità, possono rendere complicato assolvere l'onere della prova, ha introdotto un doppio binario di protezione. La consapevolezza dell'inadeguatezza delle norme nazionali ha un riverbero anche dal lato degli investimenti. Una giurisprudenza ambigua, priva di una normativa *ad hoc* sull'AI/R crea molte aspettative risarcitorie, blocca il sistema assicurativo e, di fatto, lo sviluppo tecnologico. Il doppio binario si realizza con l'adeguamento della disciplina già esistente sulla responsabilità extracontrattuale oggettiva del produttore per danno da prodotti difettosi e con l'introduzione di una norma sulla responsabilità extracontrattuale per colpa da fatti derivanti da AI/R. In particolare, la Direttiva attiene specificatamente alle azioni di responsabilità extracontrattuale, con imputabilità per colpa. La Direttiva pone due strumenti potenti nelle mani del ricorrente/danneggiato. Il primo strumento riguarda la possibilità di richiedere al giudice l'ordine di divulgazione degli elementi di prova. Si tratta di una fase preliminare giudiziale mediante cui si può ottenere dal giudice competente l'ordine di far divulgare quegli elementi di prova che sono ritenuti rilevanti in relazione all'AI/R ad alto rischio che si sospetta abbiano causato danni. Il giudice dovrebbe ordinare tale divulgazione limitatamente a ciò che serve per sostenere una domanda di risarcimento. Il secondo strumento attiene all'introduzione di una presunzione relativa che è volta a definire il nesso di causalità tra la non conformità dell'AI/R agli *standards* europei (v. Regolamento) e la condotta/il risultato dell'AI/R (o, eventualmente, la mancata produzione di un risultato da parte del sistema dell'AI/R) che ha causato il danno.

A questo punto, avendo individuato le norme euro-unitarie che più possono interessarci per fini di questo studio, proviamo a indicare di seguito dove esse possono intersecare il lavoro e la gestione dei rischi da lavoro. In particolare, si può cogliere un primo spunto nel Considerando 36 del Regolamento (secondo la versione COM/2021/206 final), in parte modificato e rinumerato come Considerando 57 nella versione approvata dal Parlamento europeo il 13 marzo 2024. Tale Considerando esprime un giudizio assai incisivo, in negativo, forse anche sproporzionato, sull'utilizzo dell'AI/R nella gestione dei lavoratori, ammonendo che per l'assunzione e la selezione delle persone, per l'adozione di decisioni in materia di promozione e cessazione dei contratti di lavoro, nonché per l'assegnazione di mansioni, per il monitoraggio o la valutazione dei lavoratori “dovrebbero essere classificati come sistemi ad alto rischio, in quanto tali sistemi possono avere un impatto significativo sul futuro di tali persone in termini di future prospettive di carriera e sostentamento”. Si insiste perché “durante tutto il processo di assunzione, nonché ai fini della valutazione e della promozione delle persone o del proseguimento dei rapporti contrattuali legati al lavoro, tali sistemi possono perpetuare modelli storici di discriminazione, ad esempio nei confronti delle donne, di talune fasce di età, delle persone con disabilità o delle persone aventi determinate origini razziali o etniche o un determinato orientamento sessuale”.

Sono tutti temi già sondati, anche in studi scientifici recenti²², su cui man mano si sta formando una dottrina articolata e plurale. Qui, invece, il punto di ricaduta della riflessione che segue è volutamente diverso, cioè più consono alla dinamica della fabbrica (nel senso di unità produttiva) che si evolve digitalmente perché si decide di mutare la struttura produttiva-organizzativa orientandola alla robotica e all'intelligenza artificiale²³ e, di conseguenza, più in linea con l'importante sistema procedimentale di anticipazione e mitigazione del rischio da AI/R nei luoghi di lavoro disegnato dal protocollo europeo del giugno del 2020, sottoscritto da organizzazioni sindacali e organizzazioni datoriali²⁴, e ripreso dall'art. 26,

²² Rinvio ai saggi pubblicati in q. Rivista negli ultimi due anni. In particolare, per i fini di questo studio, richiamo C. FALERI, *Management algoritmico e asimmetrie informative di ultima generazione*, in q. Rivista, n. 3, 2024, pp. 217 ss.; M. CORTI, *L'intelligenza artificiale nel decreto trasparenza e nella legge tedesca sull'ordinamento aziendale*, in q. Rivista, n. 29, 2023, pp. 163 ss.; U. GARGIULO, *Intelligenza Artificiale e poteri datoriali: limiti normativi e ruolo dell'autonomia collettiva*, in q. Rivista, n. 29, 2023, pp. 171 ss.; L. IMBERTI, *Intelligenza artificiale e sindacato. Chi controlla i controllori artificiali?*, in q. Rivista, n. 29, 2023, pp. 192 ss.; A. ALAIMO, *Il regolamento sull'intelligenza artificiale*, in q. Rivista, n. 25, 2023, pp. 133 ss.; L. TEBANO, *Poteri datoriali e dati biometrici nel contesto dell'AI Act*, in q. Rivista, n. 25, 2023, pp. 198 ss. Si v. anche S. CIUCCIOVINO, *La disciplina nazionale sulla utilizzazione della intelligenza artificiale nel rapporto di lavoro*, in *Lavoro, Diritti, Europa*, gennaio 2024; M. MAGNANI, *L'intelligenza artificiale e il diritto del lavoro*, in *Bollettino ADAPT*, gennaio 2024; M. PERUZZI, *Intelligenza artificiale e lavoro. Uno studio su poteri datoriali e tecniche di tutela*, Giappichelli, Torino, 2023.

²³ Si v., per esemplificare, il recente caso Saipem, nel cui contratto collettivo aziendale del 15 gennaio 2024 sono state definite alcune misure di tutela dei lavoratori nell'ambito di sistemi di intelligenza artificiale di nuova generazione. Si tratta di un costante monitoraggio attuato mediante una “soluzione tecnologica di intelligenza artificiale” che permette di prevenire i comportamenti, anche inconsapevoli, di inosservanza degli obblighi di sicurezza nei cantieri. Vengono elaborate immagini provenienti da telecamere poste nel cantiere, con tutte le cautele che la legge dispone, e, l'intelligenza artificiale, selezionando i casi di rischiosità, invia notifiche *smart* ai lavoratori in pericolo.

²⁴ Si v. il protocollo ETUC – Business Europe in https://www.etuc.org/system/files/document/file2020-06/Final%2022%2006%2020_Agreement%20on%20Digitalisation%202020.pdf - Si v. anche il Considerando 92 del



co. 7, del Regolamento (versione di marzo 2024), il quale dispone l'obbligo per il *deployer* dell'intelligenza artificiale a alto rischio di informare le rappresentanze sindacali e i lavoratori. In particolare, abbiamo segnalato che c'è un campo elettivo per svolgere questa indagine che è stato, almeno a oggi, poco investigato e che si ritiene centrale per poter rileggere in modo unitario tutti gli aspetti che sono già oggetto di analisi accademica (AI/R e esercizio dei poteri datoriali). Si tratta di un ambito scientifico molto interessante e ritenuto utile anche al fine di individuare il substrato di una nuova branca del diritto del lavoro che qui chiamiamo "RLL", *Robot Labor Law* perché muove dall'osservazione della realtà aziendale dove operano AI/R in piena interazione con le persone. Si tratta del campo di analisi riferito ai danni alla persona del lavoratore, e, indirettamente a quello delle misure di mitigazione di tali danni, tra cui i procedimenti di sicurezza sui luoghi di lavoro e il regime assicurativo per le tecnopatie (INAIL). Come anticipato sopra, la tecnologia avanzata (*Frontier AI*) viene intesa ai fini della nostra ricerca, da un lato, come strumento da cui può dipendere anche un danno alla persona del lavoratore, in ragione di una violazione delle norme sulla sicurezza, e, dall'altro, come strumento che aiuta a prevenire il danno ai lavoratori, in ragione di elementi che abilitano nuove interazioni tra macchina e persona, nuove competenze, nuove forme di intervento intelligente a protezione della persona (exp. esoscheletro, forme remotizzate di controllo e monitoraggio intelligente, forme innovative di intervento tecnologico, attuato mediante droni o altri *robot* antropomorfi, in caso di probabile rischio, anche in sostituzione dei lavoratori nell'ambito di produzioni di vetro, gomma, tessuti, chimica, farmaceutica, energia, etc.)²⁵.

Sebbene sotto il primo profilo (danno causato dall'AI/R) ci siano almeno due elementi da studiare per verificare l'adeguatezza dell'attuale assetto normativo in materia di rischio professionale e responsabilità, non si può non sottolineare che, a monte di tali elementi, c'è una premessa metodologica da ribadire, anche per essere ancora più netti nella gerarchia delle fonti che viene a crearsi in ragione della distinzione tra materie di competenza euro-unitaria, su cui il Regolamento e la Direttiva intervengono, e materie giuslavoristiche e previdenziali, al momento di competenza domestica: da una parte, c'è la norma europea che disciplina ciò che è vietato, ciò che si può fare e ciò che non si dovrebbe fare in assenza di alcuni requisiti (v. sopra), e, dall'altra, la norma domestica previdenziale assicura la protezione dal danno derivante da rischio professionale, con la conseguenza che qualora una certa attività tecnologica fosse

Regolamento della versione di marzo 2024 che dispone il diretto collegamento con il diritto dei lavoratori a essere informati e consultati, anche per il tramite delle rappresentanze sindacali, su tutto ciò che attiene all'intelligenza artificiale a alto rischio applicata nei luoghi di lavoro.

²⁵ Le realtà aziendali che abbiamo invitato a partecipare nel progetto di ricerca sono Eni, Enel, Ferrovie dello Stato, Leonardo, Hera, O-I International, Intesa San Paolo, Saipem. Tra le prime associazioni datoriali che hanno dato la disponibilità a far parte del lavoro di ricerca ci sono Confcommercio, Federdistribuzione, ABI, Confartigianato, CNA, Federfarma, Confindustria Digitale Anitec-Assinform. CGIL, CISL, UIL sono state informate del progetto.

vietata dalla norma europea, difficilmente essa potrebbe essere oggetto di assicurazione obbligatoria contro gli infortuni sul lavoro e le malattie professionali.

I due elementi che studiamo per verificare l'adeguatezza delle regole sull'assicurazione INAIL da danni derivanti da AI/R nei luoghi di lavoro sono riferiti al sistema di sicurezza sociale italiano, tenendo in considerazione che questo saggio rappresenta una prima fase di un'indagine, anche comparativa, *in fieri* e più ampia.

Il primo elemento riguarda la tutela da rischi da lavoro, la quale viene accordata dal legislatore esclusivamente ai lavoratori che sono generalmente esposti a un certo tipo di rischio lavorativo. In particolare, secondo la giurisprudenza (Cost. 2 marzo 1991, n. 100), sussiste una dissociazione di presupposti tra assicurazione obbligatoria contro gli infortuni sul lavoro e metodo statistico-assicurativo di rischio. Tale dissociazione deriva dall'art. 38, co. 2, Cost., secondo il quale il sistema di protezione assicurativa non è volto a garantire in sé il rischio di infortunio o di malattia professionale, “bensì questi eventi in quanto incidenti sulla capacità di lavoro e collegati da un nesso causale a un'attività tipicamente valutata dalla legge come meritevole di tutela”. Oggetto della tutela sono, dunque, alcune attività tipiche, al di là dall'entità concreta della relativa pericolosità. Il che determina un quadro complesso da analizzare perché in tutti i settori produttivi il legislatore nel tempo ha selezionato i lavoratori la cui attività è ritenuta, per valutazioni politiche e sociali, legate a un certo momento storico, più esposta di altre al rischio di infortunio/malattia. Le lavorazioni, almeno nel sistema italiano, sono definite pericolose mediante un rinvio doppio: da una parte, esse sono pericolose se si svolgono in base all'utilizzo di talune macchine, le cui caratteristiche sono individuabili di volta in volta dalla norma, e, dall'altra, a prescindere da tali macchine, esse sono pericolose, se rientrano in un elenco tassativo, non suscettibile di applicazione analogica.

Il secondo elemento attiene alla garanzia economica, la quale non è equivalente al danno subito. C'è una fisiologica discrasia tra entità del danno al lavoratore e prestazione economica erogabile in quanto non si tratta di un risarcimento del danno, ma di un mero indennizzo, non potendo l'assicurazione INAIL coprire tutti i danni subiti. Il che determina indirettamente la giustificazione del cd. esonero del datore di lavoro dalla responsabilità civile *ex art. 10 DPR 30 giugno 1965, n. 1124*. Tale regime esonerativo è parziale, anzi si può ritenere che esso ormai sia assai ridotto nella relativa funzione di garanzia rispetto al datore di lavoro. Essendo esso frutto di una lunga rimodulazione interpretativa giurisprudenziale, prevalentemente operata dalla Corte costituzionale²⁶, si può affermare che la tutela del lavoratore che abbia subito un infortunio sul lavoro o una malattia professionale viene oggi assegnata, nel sistema

²⁶ Tra le più importanti, si v. Cost. 11 luglio 2003, n. 233; Cost. 24 aprile 1986, n. 118; Cost. 19 giugno 1981, n. 102; Cost. 9 marzo 1967, n. 22. Si v. la ricostruzione recente di G. CORSALINI, *L'azione di regresso dell'INAIL e il significato della sua autonomia*, in *Rivista del Diritto della Sicurezza Sociale*, n. 3, 2023, pp. 617 ss.

italiano, al concorso tra regole INAIL (assicurazione obbligatoria) e regole del codice civile sulla responsabilità del datore di lavoro. Dove le prime (INAIL) cessano di operare, intervengono le seconde (codice civile): il datore di lavoro, teoricamente, non sarebbe chiamato a rispondere con riferimento alle voci di danno assorbite dall'assicurazione INAIL, ma ciò non accade quasi mai perché qualora ci sia un reato perseguibile di ufficio e una sentenza stabilisca che l'infortunio sia avvenuto per un fatto riferibile al medesimo datore o al relativo incaricato alla direzione/sorveglianza, il datore di lavoro viene chiamato a rispondere di tutti i danni arrecati al lavoratore²⁷. Inoltre, per i danni non coperti dall'assicurazione obbligatoria, i cd. danni complementari, non c'è esonero di sorta e si applica integralmente il sistema risarcitorio di cui al codice civile (Cost. 18 luglio 1991, n. 356).

Alla luce della descrizione dei due elementi, le linee di ricerca da sviluppare per alimentare un dibattito sui presupposti teorici della nuova branca del diritto del lavoro (*Robot Labor Law* – RLL) possono muovere da alcune domande da porsi. Ci sono due preliminari domande sull'adeguatezza di tale regime rispetto ai danni da AI/R sui luoghi di lavoro: (i) può il sistema (probabilmente vetusto) di classificazione della pericolosità della lavorazione essere ritenuto altresì sufficiente a garantire il lavoratore rispetto all'AI/R che opera in unità produttive avanzate? (ii) Può il sistema di imputazione di responsabilità del datore di lavoro, così come lo conosciamo noi, essere applicato anche *sic et simpliciter* nel caso di AI/R, anche nella forma di *Frontier AI*?

E ancora, nell'ambito di tali domande, ci sono ulteriori spazi di indagine: sarebbe sufficiente estendere l'elenco INAIL delle lavorazioni pericolose, assorbendo anche l'AI/R? E poi quale AI/R, anche la *Frontier AI* o solo la *Frontier AI*? E nel caso in cui l'AI/R controllasse indirettamente le macchine già considerate pericolose? O, al contrario, qualora l'AI/R controllasse, anche in relazione alle *downstream tasks*, macchine non considerate utili per definire lavorazioni pericolose o non elencate, ma in ogni caso, in ragione della gestione da parte dell'AI/R, capaci di arrecare danno? Come si deve re-interpretare la nozione di prevenzione di fronte alla *Frontier AI*, e, dunque, la responsabilità per scelte che non possono essere imputate al datore di lavoro? E, dunque, perché il datore di lavoro dovrebbe rispondere dell'operato di un'AI/R capace di auto-determinare (quasi o sempre più) integralmente la propria condotta, come nel caso della *Frontier AI*? Cosa significa sistema di prevenzione del rischio/danno da AI/R nell'ambito di

²⁷ Si tenga in considerazione che l'art. 590 c.p. estende a tutte le ipotesi di lesione colposa la regola della procedibilità a querela di parte, con l'eccezione delle lesioni gravi o gravissime commesse con violazione delle norme di prevenzione degli infortuni/malattie professionali e relative all'igiene del lavoro. Il che non restringe l'area dell'esonero del datore di lavoro. Anzi, al contrario, facendo perno sulla funzione di prevenzione dell'art. 2087 c.c. e sul d.lgs. 9 aprile 2008, n. 81, la giurisprudenza ha ritenuto che qualsiasi inosservanza dell'obbligo di sicurezza integra la fattispecie di cui all'art. 590 c.p. (si v. *ex multis* Cass. 25 agosto 1995, n. 9000; Cass. 12 febbraio 2000, n. 1579), con la conseguenza che il datore di lavoro viene generalmente chiamato a rispondere di tutti i danni cagionati al lavoratore qualora l'evento integri gli estremi di un reato perseguibile di ufficio e ci sia la violazione degli obblighi di sicurezza sul lavoro, con la configurazione ex art. 43 c.p. di una responsabilità per colpa.

un'organizzazione aziendale plasmata, controllata, monitorata, coordinata da sistemi intelligenti che si sostituiscono, in termini di terzo elemento del rapporto di lavoro, al datore di lavoro nell'esercizio dei poteri tipici di controllo, coordinamento, monitoraggio, etc.? Quali procedimenti bisogna porre in essere? Quali sono quelli più efficienti per garantire la persona del lavoratore e, contestualmente, non bloccare l'innovazione tecnologica e finanziaria che il datore di lavoro potrebbe effettuare? Quali tutele e quali procedimenti di esame congiunto si possono negoziare a livello aziendale, tenendo in considerazione anche le migliori prassi europee?

Il che, però, non sarebbe in ogni caso sufficiente perché bisogna altresì osservare la fase *in fieri* e quella *ex post* all'introduzione di una AI/R a livello di unità produttiva, ponendosi ulteriori domande: in questa prospettiva, è ancora efficiente, ai fini della mitigazione del danno, la proceduralizzazione che l'attuale regolazione sulla sicurezza sul lavoro impone²⁸, anche nell'ambito di produzioni a tecnologia avanzata o dove il lavoratore co-opera con l'AI/R? La conformità, anche certificata da soggetti terzi, a tale assetto di regole proceduralizzate della sicurezza sul lavoro può determinare in presenza di AI/R forme di esonero più efficienti per il datore di lavoro? Tra le misure di prevenzione del rischio da AI/R vi possono rientrare funzioni speciali delle istituzioni paritetiche che si occupano di sanità integrativa (fondi sanitari integrativi ex d.lgs. 30 dicembre 1992, n. 502) o di sicurezza sul lavoro (d.lgs. 9 aprile 2008, n. 81)? Tra le funzioni speciali dei fondi sanitari integrativi si possono annoverare, per indicazione della contrattazione collettiva istitutiva, forme di monitoraggio della salute psico-fisica riferibili alla tecnologia avanzata e AI/R più utilizzata a livello aziendale e/o settoriale? I dati sulla salute psico-fisica rielaborati da tali fondi sanitari integrativi potrebbero essere utilizzati per definire mediante contratto collettivo in modo più adatto le politiche sanitarie di prevenzione del rischio da lavoro, con specifica riferibilità al tipo di rischio, al settore, all'AI/R utilizzata? Per la rielaborazione di tali dati si può immaginare l'utilizzo di un libretto digitale del lavoratore/cittadino (*digital wallet*), a cui hanno accesso i fondi sanitari integrativi e altre istituzioni paritetiche, anche per far confluire informazioni su formazione in materia di sicurezza, situazioni di rischio psico-sociale, etc.? Sarebbe auspicabile predisporre un *digital wallet* per la registrazione della formazione in materia di sicurezza, in cui far confluire anche informazioni sulla prevenzione e sulla salute del lavoratore che coopera con l'AI/R? L'applicazione delle norme contrattuali

²⁸ Sul punto si v. gli studi di P. PASCUCCI, *Salute e sicurezza sul lavoro, responsabilità degli enti, modelli organizzativi e gestionali*, in *Rivista giuridica del lavoro e della previdenza sociale*, n. 4, 2021, pp. 537 ss. nonché *Sicurezza sul lavoro e cooperazione del lavoratore*, in *Giornale di diritto del lavoro e di relazioni industriali*, 2021, n. 171, pp. 421 ss. Per una ricostruzione dettagliata dei temi più rilevanti sulla responsabilità datoriale, anche in contesti tecnologicamente aziendali evoluti, si v. il recente libro monografico di M. GIOVANNONE, *Responsabilità datoriale e prospettive regolative della sicurezza sul lavoro. Una proposta di ricomposizione*, Giappichelli, Torino, 2024. Si v. alcuni spunti nei saggi di F. MARINELLI, *Verso una Fabbrica Intelligente: come l'AI invita a ripensare la tutela della salute e della sicurezza dei lavoratori*, in *Variations su Temi di Diritto del Lavoro*, n. 4, 2023, pp. 828 ss. e F. MALZANI, *Tassonomia UE e vincoli per l'impresa sostenibile nella prospettiva prevenzionistica*, in *Giornale di diritto del lavoro e di relazioni industriali*, n. 177-178, 2023, pp. 75 ss.

sui fondi di sanità integrativa che dispone questo tipo di iniziativa prevenzionale del rischio può determinare una riduzione del premio INAIL in ragione della certificata conformità a taluni protocolli sanitari di anticipazione del rischio psico-fisico da AI/R? E, ancora, sarebbe possibile prevedere una riduzione del premio INAIL per i datori di lavoro che investono di più e meglio in tecnologia avanzata certificata volta a mitigare il rischio e/o a introdurre forme di prevenzione mirata del rischio mediante l'intervento di fondi di sanità integrativa?

4. Rischio, danno alla persona e norma statunitense. Cenni

A fine 2023 è stato emanato dal Presidente degli Stati Uniti d'America un provvedimento che regola specificatamente la *Frontier AI*. Si tratta dell'*Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, di seguito anche "EO", il quale è parte di una strategia più ampia composta dal *Blueprint for an AI Bill of Rights*, dall'*AI Risk Management Framework* e dall'istituzione del *National AI Research Resource*²⁹.

L'EO impone alle imprese che si occupano di intelligenza artificiale di svolgere il cd. *AI Red Teaming*, il quale è definito in termini procedurali interni, di livello aziendale, come una sperimentazione strutturata sulle possibilità di vulnerabilità dell'AI/R da cui possono derivare rischi di vario genere e danni alle persone umane ("a structured testing effort to find flaws and vulnerabilities in an AI system, often in a controlled environment and in collaboration with developers of AI. Artificial Intelligence red-teaming is most often performed by dedicated "red teams" that adopt adversarial methods to identify flaws and vulnerabilities, such as harmful or discriminatory outputs from an AI system, unforeseen or undesirable system behaviors, limitations, or potential risks associated with the misuse of the system"). Sono obbligate quelle imprese che elaborano sistemi di intelligenza artificiale di nuova generazione, qui definiti anche *Frontier AI*, o, altresì denominati "*Dual-Use Foundational Model*". Il che implica una indiretta selezione dei soggetti vincolati a tale obbligo in relazione alla capacità esprimibile dalla macchina intelligente. Per la norma statunitense è *Dual-Use Foundational Model* il sistema istruito in relazione a una ampia gamma di dati, capace di auto-supervisionarsi, in grado di gestire miliardi di parametri, applicabile trasversalmente, in molti contesti, capace di svolgere compiti da cui possono derivare rischi gravi per la sicurezza in generale, per quella nazionale e economica, per quella relativa alla sanità dei cittadini, per combinazioni di queste, come l'incidenza sulle barriere protettive che concernono l'uso di elementi chimici, biologici, nucleari, attacchi cyber, etc. o la deviazione dal controllo umano. In

²⁹ Si v. la ricostruzione di R. CALO, *Artificial Intelligence Policy: A Primer and Roadmap*, in 51 *U.C. Davis L. Rev.*, n. 51, 2017, pp. 399 ss. e, più recentemente, di M. E. KAMINSKI, *Regulating the Risks of AI*, in *Boston University Law Review*, n. 103, 2023, pp. 1347 ss. Si v. per i temi giuslavoristici che intrecciano la regolazione dell'intelligenza artificiale negli Stati Uniti d'America le analisi di B. ROGERS, *Data and Democracy at Work. Advanced Information Technologies, Labor Law, and the New Working Class*, MIT, Cambridge, 2023; O. LOBEL, *The Equality Machine: Harnessing Digital Technology for a Brighter, More Inclusive Future*, Public Affairs, New York, 2022.

attesa di un consolidamento della definizione normativa di *Dual-Use Foundational Model*, l'EO rileva, in via preliminare, assegnando ai competenti segretari di stato l'ulteriore aggiornamento, una serie di dettagli tecnici per selezionare cosa è già oggi *Dual-Use Foundational Model*. In particolare, nell'art. 4.2, si individuano alcuni elementi che attengono, da una parte, alla potenza del modello di AI/R (*any model that was trained using a quantity of computing power greater than 1026 integer or floating-point operations, or using primarily biological sequence data and using a quantity of computing power greater than 1023 integer or floating-point operations*), e, dall'altra, alla capacità dell'infrastruttura digitale (*any computing cluster that has a set of machines physically co-located in a single datacenter, transitively connected by data center networking of over 100 Gbit/s, and having a theoretical maximum computing capacity of 1020 integer or floating-point operations per second for training AI*).

Inoltre, l'EO impone ai fornitori di infrastrutture di servizio, come *Amazon, Google, Microsoft, etc.*, uno specifico obbligo che consiste nel monitorare e riportare alle autorità pubbliche le condotte di cittadini stranieri che istruiscono sistemi di intelligenza artificiale con le caratteristiche di potenza di cui sopra anche per eventuali attacchi *cyber* alla sicurezza del paese (cd. "*malicious cyber-enabled activity*"). Il che significa aver fatto una scelta diversa da quella europea: nell'ordinamento statunitense si svolge un monitoraggio di tutte le possibili applicazioni di AI/R, anche di quelle di nuova generazione, mediante il sistema dei grandi *clouds*. La parte positiva di questa scelta viene compensata dalle molte criticità, le quali sono riferibili a possibili abusi da parte dei fornitori delle infrastrutture di *cloud* o alla capacità di soggetti di effettuare vere indagini sulle possibili *malicious cyber-enabled activities*.

Abbiamo segnalato sopra che il genotipo che si può costruire per avviare la comparazione tra norma europea e norma statunitense si pone in relazione a tre problemi: (i) quale rischio e quali danni quell'ordinamento giuridico intende selezionare ai fini della tutela da predisporre in relazione alla capacità della *Frontier AI* di interagire con il lavoratore? (ii) Quali misure di mitigazione e prevenzione quell'ordinamento giuridico intende introdurre per far fronte al bisogno di tutela correlato a quei rischi e a quei danni? (iii) Quali istituzioni quell'ordinamento giuridico intende porre per attuare tale tutela, anche in termini di esigibilità individuale e collettiva?

Si comprende che la norma statunitense ha una più ampia portata geo-politica di quella europea, avendo essa impostato la definizione di *Frontier AI* sulla base di una serie di criteri che attengono alla potenza del sistema informatico e di quello infrastrutturale che lo sostiene³⁰. Il che può avere riflessi anche nel campo della presente indagine giuslavoristica (l'AI/R che determina o previene danni alla persona del lavoratore nelle unità produttive avanzate). In particolare, nel sistema nordamericano di tutela delle condizioni di lavoro si intrecciano almeno tre elementi, i quali hanno carattere preventivo, di mitigazione e assicurativo.

³⁰ Si v. i documenti pubblicati qui - <https://ai.gov> e la ricostruzione effettuata da alcuni studiosi della *Stanford University*, in particolare la direttrice del centro di ricerca sull'intelligenza artificiale (*HAI*), prof.ssa FEI-FEI LI, *The World I see. Curiosity, Exploration, and Discovery at the Dawn of AI*, Flatiron, USA, 2023.

Il primo elemento attiene al contenzioso individuale che ciascun lavoratore danneggiato può avviare per essere risarcito del danno subito (*Torts Suits*)³¹. Il contenzioso segue le regole generali, con l'effetto che i rimedi punitivi e il risarcimento segue l'impostazione nordamericana, producendo una certa deterrenza indiretta rispetto alla violazione di norme di sicurezza. Il secondo elemento riguarda l'erogazione di una prestazione economica e/o in servizi medici a favore del lavoratore danneggiato (*Workers' Compensation*). Si tratta di uno strumento previdenziale-assicurativo, attuato mediante il pagamento di premi a carico del datore di lavoro, il quale, in caso di infortunio o malattia professionale, viene attivato a favore del lavoratore³². Per accedere a tale sistema si dovrà indagare di volta in volta se l'incidente sul lavoro o la malattia possa determinare o meno l'erogazione della prestazione³³. Il che prevede l'effettuazione di un *test* preliminare che si basa sull'analisi dei fatti (è davvero un incidente sul lavoro? L'incidente è avvenuto “*in the course of the employment*” o “*arising out of the employment*”? In cosa consiste il danno al lavoratore?). La nozione di incidente/malattia è generalmente aperta, e caso per caso si svolge un *test* giurisprudenziale di indagine³⁴. L'applicazione del regime del *Workers' Compensation* teoricamente crea una forma di esonero del datore di lavoro da altre responsabilità e richieste di risarcimento. Ci sono, però, come nei sistemi europei, una serie di situazioni che determinano il diritto del lavoratore a agire in ogni caso per aggiungere alle prestazioni previdenziali-assicurative anche il risarcimento del danno ulteriore³⁵. Il terzo elemento è relativo alla struttura della prevenzione e della sicurezza sul lavoro, strutturata secondo il modello dell'anticipazione del rischio e della gestione degli effetti dei rischi mediante alcune istituzioni che sorvegliano l'adempimento corretto degli obblighi di protezione a carico del datore di lavoro (*Safety and Health Laws*)³⁶.

Dal nostro punto di vista, le domande già poste sopra, utili a avviare le linee di ricerca sulla RLL, in relazione all'ordinamento italiano e in parte a quello europeo, si possono riproporre, con alcune differenze, anche per il sistema nordamericano. Tali domande possono essere arricchite dalle seguenti: sono adeguate o meno le misure di prevenzione del rischio psico-sociale derivante da AI/R che opera a livello di unità produttiva? Il sistema nordamericano provvede a erogare risorse economiche sufficienti

³¹ Si v. uno dei casi più importanti da cui muove il sistema del contenzioso individuale - *Farewell v. Boston & Worcester Rail Road Corp.* *Supreme Judicial Court of Massachusetts* 45 Mass. (4 Met) 49 (1842).

³² Si v. R. EPSTEIN, *The Historical Origins and Economic Structure of Workers' Compensation Law*, in *Georgia Law Review*, n. 16, 1982, pp. 775 ss.

³³ L'amministrazione del *Workers' Compensation* varia da Stato a Stato. Alcuni Stati hanno l'amministrazione simile a quella previdenziale, con una centralizzazione delle attività in una agenzia. Altri Stati hanno meccanismi simili a assicurazioni private.

³⁴ Si v. il caso *Matthews v. R.T. Allen & Sons*, *Suprem Judicial Court of Maine*, 266 A.2d 240 (1970).

³⁵ Si v. il caso *Millison v. E.I. du Pont de Nemours & Co.*, *Supreme Court of New Jersey*, 501 A.2d 505 (1985).

³⁶ Si rinvia alla disciplina dell'agenzia OSHA descritta e riportata qui - <https://www.osha.gov> Per la riflessione più teorica sulla funzione della sicurezza sul lavoro nel sistema statunitense, si v. questo studio A. P. BARTEL, L. GLENN THOMAS, *Predation Trough Regulation: The Wage and Profit Effects of the Occupational Safety and Health Administration and the Environmental Protection Agency*, in *Journal of Law & Economics*, n. 30, 1987, pp. 239 ss.

per il risarcimento del danno e la prevenzione? È equo tale sistema con riferimento alle varie tipologie di contratto individuale di lavoro, ai lavoratori sindacalizzati e non, migranti, etc.? È più o meno efficiente un sistema aperto di definizione di incidente/malattia professionale, non pre-definito mediante elenchi o liste, con riferimento all'AI/R che opera a livello di unità produttiva?

5. Conclusioni. Per impostare i contenuti di una nuova branca della disciplina giuslavoristica (RLL - Robot Labor Law)

Qualunque riflessione non può non muovere da una fotografia dell'intero sistema che si sta evolvendo, sia tecnicamente che normativamente, a livello di ordinamenti occidentali. È ciò che nel gergo di qualunque attività rischiosa può essere altresì chiamata *Vantasner Danger Meridian*, una specie di linea che demarca, segna e cronometra la reale rischiosità. Il che allontana le tecnofobie allarmistiche e avvicina il più auspicabile atteggiamento pragmatico del regolatore.

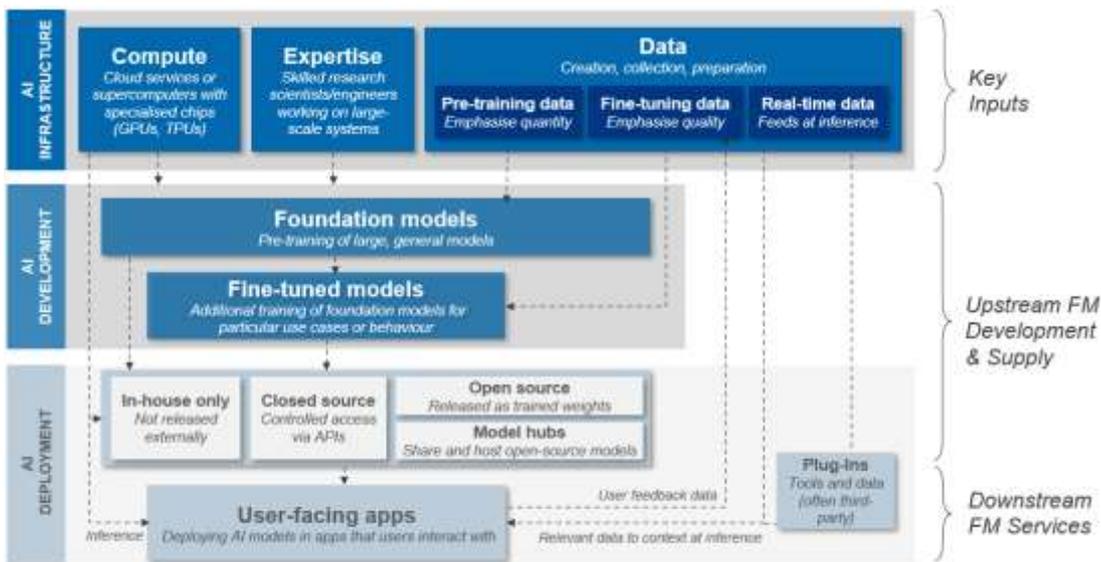
Abbiamo compreso che l'AI/R, gestita dalla *Frontier AI*, ci impone una regolazione (transnazionale/nazionale) che non può non seguire l'intero ciclo di vita della macchina intelligente, dalla fase di sviluppo iniziale sino alla sperimentazione e, poi, all'introduzione nel mercato, con verifiche periodiche degli esiti, positivi o negativi che siano, anche in termini di danno alla persona del lavoratore. Più avanzata è l'intelligenza artificiale, come nel caso della *Frontier AI*, più rischi ci sono. E ciò diventa ancora più vero per le forme di AI/R che cooperano, interagiscono, coesistono con i lavoratori nei contesti professionali.

I rischi della *Frontier AI* possono derivare da tre fattori che la regolazione, anche quella giuslavoristica, deve intercettare in qualche modo³⁷. Il primo fattore attiene ai problemi inaspettati, generalmente derivanti dal fatto che la *Frontier AI* presenta l'effettiva capacità di auto-determinarsi proprio nella fase di distribuzione nel mercato e nello svolgimento di compiti per e con gli utilizzatori finali. Il secondo fattore riguarda il problema della sicurezza nella fase di distribuzione. Gli utilizzatori finali potrebbero richiedere alla *Frontier AI* uno sviluppo ulteriore, non volto a fini legittimi, nel senso di giuridicamente non giustificabili, i quali non sono prevedibili del tutto (v. sopra) e potrebbero moltiplicare la propria portata dannosa. Il terzo fattore è relativo al problema della proliferazione, visto che la *Frontier AI* è generalmente *open-sourced*, permettendo a chiunque di sviluppare modelli ulteriori, anche aventi finalità non

³⁷ Qui si segue l'impostazione che è stata elaborata dal gruppo di studiosi e esperti, il cui lavoro è diventato oggetto di riflessione e punto di riferimento per le politiche del diritto da parte della conferenza intergovernativa che si è occupata a Londra nel 2023 della *Frontier AI*, da parte dell'Unione Europea e da parte degli Stati Uniti d'America. Si v. M. ANDERLJUNG, J. BARNHART, A. KORINEK, J. LEUNG, C. O'KEEFE, J. WHITTLESTONE, S. AVIN, M. BRUNDAGE, J. BULLOCK, D. CASS-BEGGS, B. CHANG, T. COLLINS, T. FIST, G. HADFIELD, A. HAYES, L. HO, S. HOOKER, E. HORVITZ, N. KOLT, J. SCHUETT, Y. SHAVIT, D. SIDDARTH, R. TRAGER, K. WOLF, *Frontier AI Regulation: Managing Emerging Risks To Public Safety*, 2023, in <https://arxiv.org/abs/2307.03718>

giuridicamente giustificabili. Spesso si tratta di furti di modelli o attacchi cyber-tecnologici che permettono di accedere a schemi che hanno un certo valore strategico.

Lo schema che può essere utile a comprendere tali fattori è il seguente ed è tratto dalla documentazione dell'*AI Safety Summit* di Londra, 2023:



Per far fronte a tali sfide la regolazione, anche quella giuslavoristica, sarebbe chiamata a impostare (i) *standard* internazionali di sicurezza³⁸, (ii) modelli esigibili di conformità a tali *standards*, (iii) indici e schemi di trasparenza e informazione, (iii) un modello assicurativo contro il rischio da lavoro derivante da *Frontier AI* di matrice macro-regionale, non più solo domestico/nazionale.

Muoviamo dagli *standards* internazionali, i quali dovrebbero essere oggetto di una forma di regolazione convenzionale o di trattato, sottoscritta almeno tra gli ordinamenti giuridici più rilevanti dell'occidente, tra cui certamente Stati Uniti d'America, UE e Gran Bretagna. Gli *standards* dovrebbero essere oggetto di studi, di sperimentazione, di verifica dei rischi, di raccolta dei dati e di confronto accademico, sociale, etc. Bisognerebbe introdurre schemi di *auditing* mirati alla verifica degli *standards*. Il che permetterebbe di costruire *standards* che sono elaborati mediante la tecnica del *risk assesment*, con la ponderazione di ciò che è capacità effettiva di pericolo derivante dall'AI/R (*Frontier AI*), anche a livello di unità produttiva, e della relativa controllabilità. La valutazione da parte di soggetti terzi, specializzati, *auditors* o simili, di fatto creerebbe una maggiore credibilità basata su ciò che si rende noto all'esterno, anche con l'ausilio di tali

³⁸ Negli Stati Uniti d'America, il *National Institute for Standards and Technology* ha creato l'*AI Risk Management Framework*. La *National Telecommunication and Information Agency* ha avviato un percorso per introdurre politiche assicurative collegate all'intelligenza artificiale. In Gran Bretagna è stato istituito l'*AI Standards Hub*. L'Unione europea ha chiesto alle agenzie CEN and CENELEC di sviluppare modelli standardizzati di sicurezza sull'intelligenza artificiale.

terzi specializzati. Di qui muoverebbe la costruzione di protocolli standardizzati sulle modalità di distribuzione dell'AI/R (*Frontier AI*) e sulle tutele da introdurre nonché sui processi di revisione periodica dei rischi e delle misure da adottare.

Ma tutto ciò probabilmente non sarebbe in ogni caso sufficiente. Si dovrebbero rendere esigibili i modelli di conformità, bilanciandone l'impatto sullo sviluppo potenziale tecnologico-economico. Ciò è la vera sfida per un legislatore. Si dovrebbe promuovere, anche a livello transnazionale, una via di adesione volontaria a tali *standards*, con la creazione di codici di condotta (*self-regulation*), certificazioni, rilascio di abilitazione/patente speciale preventiva per poter creare e poi distribuire, anche per le collaborazioni persona/macchina, sistemi di *Frontier AI*, autorità di vigilanza effettivamente attiva nelle ispezioni, etc. Il che permetterebbe di attivare nelle unità produttive a alta tecnologia sistemi di controllo incrociato, nuovi procedimenti di sicurezza sul lavoro collegati alla formazione professionale certificata per ciascun lavoratore che collabora con l'AI/R, patentini speciali per tali lavoratori, etc. Di qui muoverebbe, infine, la possibilità di ipotizzare un sistema di assicurazione obbligatoria, almeno macro-regionale, europeo, riferibile a ciò che in Italia è l'INAIL, per i rischi da AI/R sul lavoro. Chi studia le materie previdenziali sa che maggiore è l'ambito di applicazione, ordinamentale e geografico, di una forma assicurativa/previdenziale, maggiori saranno i benefici per tutti.

L'imponderabilità di tale impostazione riguarda specificatamente l'incidenza sulla norma che previene il danno e il rischio al lavoratore nonché permette di risarcire il danno nell'ambito dei sistemi avanzati di AI/R come la *Frontier AI*. È un'imponderabilità che deriva dal fatto che non si conosce ancora bene la combinazione più appropriata tra la prevenzione a livello di unità produttiva e il risarcimento dell'ulteriore danno, in caso di esercizio dei poteri datoriali da parte della *Frontier AI* o qualunque forma di intelligenza artificiale anche più evoluta.

Il che determina una riflessione, basata anche su studi scientifici interdisciplinari e transnazionali che sono riferiti alla realtà, non alla mera idea, come quello che si intende svolgere, con riferimento al fatto che non si può non verificare, da una parte, l'adeguatezza dei regimi previdenziali-assicurativi per incidenti sul lavoro e malattie professionali rispetto all'intelligenza artificiale di nuova generazione, adeguatezza da misurare rispetto all'obbligo contributivo, alle prestazioni e al regime di esonero di responsabilità del datore di lavoro, e, dall'altra, l'efficienza dei sistemi di sicurezza sul lavoro, essendo essi tutti o quasi tutti impostati sul passato, non sul futuro, cioè per una macchina assoggettata (con variabilità enorme) alle indicazioni della persona del lavoratore, non per una macchina, *Frontier AI*, che coordina, impone, controlla, dirige anche le prestazioni del lavoratore.

E tutto ciò rappresenta la sfida più importante per la comunità scientifica, per il decisore politico e, in definitiva, per chi opera nel sistema delle relazioni industriali, sindacati e organizzazioni datoriali. Si tratta



di una direzione da intuire e, poi, decidere di seguire, sapendo che la direzione è sempre più importante della velocità perché rappresenta il tutto sulla parte.