ISSN 1826-3534

federalismi.it
RIVISTA DI DIRITTO PUBBLICO ITALIANO, COMPARATO, EUROPEO

FOCUS LAVORO, PERSONA, TECNOLOGIA

21 MAGGIO 2025

Diritti fondamentali, rischio, autoreferenzialità, obblighi datoriali (incerti) nell'*AI Act*

di Gianfranco Peluso
Dottorando di ricerca in Intelligenza artificiale
Università degli Studi di Napoli "Federico II"

Diritti fondamentali, rischio, autoreferenzialità, obblighi datoriali (incerti) nell'*AI Act*[*]

di Gianfranco Peluso
Dottorando di ricerca in Intelligenza artificiale
Università degli Studi di Napoli "Federico II"

Abstract [It]: Il saggio esplora la tecnica normativa dell'approccio basato sul rischio nell'*AI Act*, evidenziando come il legislatore europeo non abbia effettuato un chiaro bilanciamento tra i confliggenti interessi. Dopo aver analizzato le diverse declinazioni del concetto di rischio presenti nel regolamento *AI* (a partire dalla generale definizione di rischio come «rischio di danno»), l'a. sostiene che, facendo leva sull'obbligo di monitoraggio sul funzionamento del sistema del *deployer*/datore di lavoro a tutela dei diritti fondamentali della persona, l'approccio regolamentare possa essere interpretato in chiave *fundamental right-based*. Ne consegue che alla valutazione di conformità del sistema di IA non corrisponde un'indistinta accettazione delle conseguenze pregiudizievoli per i diritti fondamentali dei lavoratori.

Title: Fundamental Rights, Risk, Self-Reference, (Uncertain) Employer Obligations in the AI Act
Abstract [En]: The essay explores the regulatory technique of the risk-based approach in the AI Act, highlighting how the European legislator has not established a clear balance between conflicting interests. After analysing the various iterations of the concept of risk in the AI regulation (starting from the general definition of risk as «risk of harm»), the author argues that, by leveraging the monitoring obligation of the deployer/employer to protect fundamental rights of the individual, the regulatory approach can be interpreted through a fundamental right-based perspective. It follows that the conformity assessment of an AI system does not correspond to an indiscriminate acceptance of detrimental consequences for workers' fundamental rights.

Parole chiave: Approccio basato sul rischio; *AI Act*; diritti fondamentali; bilanciamento; obblighi datore di lavoro.
Keywords: Risk-based approach; AI Act; fundamental rights; balancing; employer's obligations.

1. Profili definitori del rischio

Nella società moderna, definita dal sociologo tedesco Ulrich Beck «*Risk Society*»[1], l'esponenziale sviluppo tecnologico e i pericoli correlati al suo impiego hanno condotto all'elaborazione di soluzioni normative incentrate sul concetto di rischio e sulla sua valutazione.

Come ancora ricordato di recente, in ambito giuslavoristico il concetto di rischio è tradizionalmente legato all'incolumità psico-fisica del lavoratore (art. 2087 c.c.), all'insicurezza dello stesso lavoro e alla dignità

[*] Articolo sottoposto a referaggio.
[1] U. BECK, *Risk Society. Towards a New Modernity*, SAGE publications, New York, 1992.

della persona che lavora (art. 41 Cost.)[2]. Il rischio e, segnatamente, la sua preventiva valutazione contraddistinguono la disciplina in materia di salute e sicurezza sul lavoro (dir. 89/391CEE[3] e d.lgs. n. 81/2008[4]) secondo la logica della prevenzione primaria che prevede l'eliminazione dei rischi evitabili e la successiva valutazione dei soli rischi inevitabili[5]; e pure il regolamento generale sulla protezione dei dati[6] si contraddistingue per l'adozione di una tecnica regolatoria basata sul rischio (*infra* par. 3).

Nel regolamento europeo sull'intelligenza artificiale (reg. UE 2024/1689 del 13 giugno 2024, c.d. *AI Act*) – atto che, come è noto, rinviene la sua base giuridica negli artt. 16 e 114 TFUE ma che rileva in ambito lavoristico per le implicazioni che l'intelligenza artificiale (in seguito IA) e la sua regolazione comportano nel mondo del lavoro – il concetto di rischio assume connotati peculiari[7]. Attraverso la tecnica normativa del c.d. *approccio basato sul rischio*, la quantificazione del rischio – le cui variabili di calcolo riguardano la

[2] Così L. ZOPPOLI, *Il diritto del lavoro dopo l'avvento dell'intelligenza artificiale: aggiornamento o stravolgimento? Qualche (utile) appunto*, in *WP CSDLE "Massimo D'Antona".IT*, 489/2024, p. 12.

[3] Sin dalla direttiva 89/391/CEE del 12 giugno 1989, infatti, si segna «il superamento dell'impostazione tecnico-oggettiva della precedente normativa» e si sposta «il focus delle tecniche di prevenzione dalla prevenzione "oggettiva", alla prevenzione "soggettiva", fondata sulla maggiore considerazione del rapporto tra il lavoratore, l'ambiente di lavoro ed i fattori di rischio», così G. NATULLO, *Il quadro normativo dal Codice civile al Codice della sicurezza sul lavoro. Dalla Massima Sicurezza possibile alla Massima Sicurezza effettivamente applicata?*, in G. NATULLO (a cura di), *Salute e sicurezza sul lavoro*, UTET Giuridica, San Mauro Torinese, 2015, pp. 9 e 10.

[4] In particolare, ai sensi dell'art. 2, co. 1, lett. *s*, del d.lgs. n. 81/2008 per rischio si intende la «probabilità di raggiungimento del livello potenziale di danno nelle condizioni di impiego o di esposizione ad un determinato fattore o agente oppure alla loro combinazione». Individua un parallelismo con la normativa in materia di IA S. CAIROLI, *Intelligenza artificiale e sicurezza sul lavoro: uno sguardo oltre la siepe*, in *Diritto della Sicurezza sul Lavoro*, n. 2, 2024, p. 44: «L'architettura normativa dell'AI Act si presenta come una regolazione di natura "risk-based", ossia espressione di un approccio basato sul rischio, analogamente ai principi e alla filosofia generale, basati sulla preventiva valutazione del rischio, che ormai contrassegnano la disciplina comunitaria e interna in materia di salute e sicurezza sul lavoro». Sul potenziale delle strumentazioni di IA per la prevenzione e la mitigazione dei rischi v. T. TREU, M. FAIOLI, *Introduzione a una ricerca sulla sicurezza e sui rischi del lavoro: quadro normativo e prassi internazionali*, in q. *Rivista*, n. 6, 2025, p. 2 ss.; M. FAIOLI, *Adeguatezza ex art. 2086 c.c. e obbligo di introdurre tecnologia avanzata per mitigare i rischi da lavoro*, in q. *Rivista*, n. 6, 2025, p. 143 ss.

[5] Sul punto, tra i contributi più recenti, v. almeno: L. D'ARCANGELO, *Robotica e lavoro. Prime osservazioni in tema di sicurezza (delle macchine e) dei lavoratori*, in q. *Rivista*, n. 6, 2025, p. 83 ss.; M.G. ELMO, *Sistemi IA e rischi per la salute e la sicurezza dei lavoratori: riflessioni a margine della regolamentazione europea*, in *AmbienteDiritto*, 2024, p. 571 ss.; C. LAZZARI, P. PASCUCCI, *Sistemi di IA, salute e sicurezza sul lavoro: una sfida al modello di prevenzione aziendale, fra responsabilità e opportunità,* in *Rivista giuridica del lavoro e della previdenza sociale*, n. 4, 2024, I, p. 587 ss.; P. PASCUCCI, *Le nuove coordinate del sistema prevenzionistico*, in *Diritto della Sicurezza sul Lavoro*, n. 2, 2024, p. 37 ss.; M. PERUZZI, *Sistemi automatizzati e tutela della salute e sicurezza sul lavoro,* in *Diritto della Sicurezza sul Lavoro*, 2024, n. 2, p. 78 ss.

[6] Reg. (UE) 2016/679, c.d. GDPR.

[7] Su rischio tecnologico e tecniche di tutela, cfr., *ex multis*, P. LOI, *Il rischio proporzionato nella proposta di regolamento sull'IA e i suoi effetti nel rapporto di lavoro*, in q. *Rivista*, n. 4, 2023, p. 239 ss.; A. MANTELERO, M. PERUZZI, *L'*AI Act *e la gestione del rischio nel sistema integrato delle fonti*, in *Rivista giuridica del lavoro e della previdenza sociale*, n. 4, 2024, I, p. 517 ss.; M. PERUZZI, *Intelligenza artificiale e lavoro. Uno studio sui poteri datoriali e tecniche di tutela*, Giappichelli, Torino, 2023; F.V. PONTE, *Intelligenza artificiale e lavoro. Organizzazione algoritmica, profili gestionali, effetti sostitutivi*, Giappichelli, Torino, 2024, p. 30 ss.; G. SMORTO, *Distribuzione del rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale. Una riflessione critica*, in *Il Foro italiano*, n. 5, 2024, p. 208 ss. Sui concreti rischi legati all'innesto dell'intelligenza artificiale (e dei sistemi algoritmici in generale) all'interno delle organizzazioni di lavoro, v. V. MANDINAUD, A. PONCE DEL CASTILLO, *AI systems, risks and working conditions*, in A. PONCE DEL CASTILLO (a cura di), *Artificial intelligence, labour and society*, ETUI, Brussels, 2024, p. 237 ss.; M. PAPE, *Addressing AI risks in the workplace. Workers and algorithms, in europarl.europa.eu, European Parliamentary ResearchService,* 2024, p. 3.

salute, la sicurezza e i diritti fondamentali – viene elevata a fattore di differenziazione di vincoli e tutele[8]. È come se il legislatore, considerata la volatilità delle questioni in gioco, preferisca ricorrere a tecniche adattive calibrate a suoi obiettivi nel tentativo di evitare vincoli eccessivi ed effetti paralizzanti su settori strategici ad alta intensità di capitale[9], superando l'approccio binario (presenza o assenza del rischio, con correlata adozione di misure di prevenzione). Ricorre, così, a complessi modelli regolatori nei quali il fattore rischio viene scomposto in più livelli, ai quali corrispondono regimi normativi differenti che si applicano a definite categorie di sistemi di IA. Come si legge nel considerando 26, al fine di «introdurre un insieme proporzionato ed efficace di regole vincolanti per i sistemi di IA», ci si avvale di un approccio basato sul rischio che mira ad «adattare la tipologia e il contenuto di dette regole all'intensità e alla portata dei rischi che possono essere generati dai sistemi di IA». Questo è dunque – stando alla definizione *dell'AI Act* – l'approccio basato sul rischio: differenziazione della disciplina dei "prodotti intelligenti", *rectius* prodotti che imitano l'intelligenza umana[10], in base al tipo di rischio, individuabile anche verificando l'uso o l'ambito di applicazione del prodotto.

Il presente scritto, sulla scia degli studi che hanno esaminato quest'approccio in chiave giuslavoristica[11], si propone di insistere sul rapporto tra la tecnica normativa e i diritti fondamentali dei lavoratori[12]. Questi diritti, per il "rango" che occupano nel quadro delle tutele ordinamentali, si caratterizzano per una significativa resistenza a logiche di contemperamento meramente economicistico e sollecitano una riflessione sull'opportunità di ricorrere alla suddetta tecnica normativa per garantirne la tutela.

Il tentativo di analizzare l'approccio basato sul rischio attraverso il prisma della tutela dei diritti fondamentali, nello specifico contesto dalla regolamentazione in materia di IA, muoverà da alcune considerazioni sul rapporto tra l'approccio *risk-based* e il principio di precauzione (*infra* par. 2).

---

[8] Questa nuova concezione del rischio viene evidenziata da L. ZOPPOLI, *Il diritto del lavoro dopo l'avvento dell'intelligenza artificiale, cit.*, p. 12, che richiama il Rapporto Aspen Institute Italia-Osservatorio permanente sull'adozione e l'integrazione della intelligenza artificiale, *Rapporto intelligenza artificiale 2024*, p. 119, e il *Pact for the Future, Global Digital Compact and Declaration on Future Generations* dell'Onu del settembre 2024.

[9] Così A. ALOISI, V. DE STEFANO, *Between risk mitigation and labour rights enforcement: Assessing the transatlantic race to govern AI-driven decision-making through a comparative lens*, *in European Labour Law Journal*, n. 14(2), 2023, p. 293, i quali fanno riferimento a C. QUELLE, *Enhancing Compliance under the General Data Protection Regulation: The Risky Upshot of the Accountability- and Risk-based Approach*, in *European Journal of Risk Regulation,* 2018, 9, p. 502.

[10] Come è noto, l'equivoco lessicale affonda le radici nell'*Imitation Game* di A. TURING (*Computing Machinery and Intelligence*, in *Mind*, n. 236, 1950, pp. 433-460), al quale si ricorre per valutare l'intelligenza della macchina: questa può considerarsi intelligente se un giudice umano non è in grado di distinguere affidabilmente il testo generato dalla stessa da quello elaborato dal pensiero umano.

[11] Cfr., per tutti, M. BARBERA, *"La nave deve navigare". Rischio e responsabilità al tempo dell'impresa digitale, in Labour & Law Issues*, n. 2, 2023, p. 2 ss.; P. LOI, *Il rischio proporzionato, cit.*, p. 239 ss.; M. PERUZZI, *IA e obblighi datoriali di tutela del lavoratore: necessità e declinazioni dell'approccio* risk-based, in q. *Rivista*, n. 30, 2024, p. 225 ss.

[12] Sul punto v. M. DELFINO, *Fundamental rights in the age of Robotics and Artificial Intelligence*, in A. PERULLI, T. TREU (eds), *Labour Law in the Mirror: Categories, Values, Interlocutors*, Wolters Kluwer, Alphen aan den Rijn, 2025, p. 126 ss. Più in generale sul tema, cfr. R. DEL PUNTA, *I diritti fondamentali e la trasformazione del diritto del lavoro*, in *WP CSDLE "Massimo D'Antona".IT*, 333/2017.

Come si vedrà, un elemento peculiare della recente regolamentazione europea riguarda, anzitutto, la stessa definizione di *rischio* in riferimento ai sistemi di IA, poi la definizione di *rischio sistemico* specificamente legato ai modelli di IA per finalità generali (tra i quali rientrano i modelli di IA generativa).

La specificità dell'*AI Act* risalta considerando la diversa visuale seguita dal regolamento generale sulla protezione dei dati, che pure adotta una tecnica regolatoria basata sul rischio (*infra* par. 3).

Nel GDPR, infatti, seguendo la concezione "*right-based*"[13], il *rischio di violazione dei diritti* viene tenuto ben distinto dal *rischio di danno*[14] e l'analisi del rischio si fonda sui diritti e sulla loro rilevanza[15]. Al contrario, nel regolamento europeo sull'IA, all'art. 3, pt. 2, il *rischio* viene definito come «la combinazione della probabilità del verificarsi di un *danno* e la *gravità del danno stesso*»[16]. Parimenti, in relazione alla specifica ipotesi di *rischio sistemico*, all'art. 3, pt. 65, si fa riferimento a «*effetti negativi effettivi* o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso»[17]. Sicché, nel regolamento europeo sull'IA il rischio risulta strettamente legato alla nozione di danno, secondo una concezione che potrebbe apparire "*harm-based*", ossia legata a un concreto pregiudizio.

Come si proverà a dimostrare, questa lettura può essere superata mediante un'interpretazione sistematica che, considerando l'*AI Act* nella sua interezza, ne colga la coerenza con il quadro assiologico della Carta dei diritti fondamentali dell'Unione europea, riconducendolo così a un modello "*right-based*" o, se si preferisce, "*fundamental right-based*".

2. Oltre il tradizionale principio di precauzione

Come accennato, l'approccio basato sul rischio si distingue essenzialmente per il ricorso a una logica scalare, ossia una differenziazione che richiede una responsabilizzazione dei soggetti, tanto pubblici quanto privati, rispetto ai rischi e agli effetti collaterali delle loro attività.

[13] L. MOEREL, *Big Data Protection. How to Make the D raft EU Regulation on Data Protection Future Proof,* 2014, Tilburg University, Tilburg, p. 25 ss.

[14] Cfr. Article 29 Data Protection Working Party, *Statement on the role of a risk-based approach in data protection legal frameworks,* WP218, 2014, p. 4, ove si precisa che: «The risk-based approach goes beyond a narrow "harm-based-approach" that concentrates only on damage and should take into consideration every potential as well as actual adverse effect, assessed on a very wide scale ranging from an impact on the person concerned by the processing in question to a general societal impact (e.g. loss of social trust)».

[15] A. MANTELERO, *Il nuovo approccio della valutazione del rischio nella sicurezza dei dati. Valutazione dell'impatto e consultazione preventiva (Artt., 32-39)*, in G. FINOCCHIARO (opera diretta da), *Il nuovo Regolamento europeo sulla privacy e sulla protezione dei dati personali*, Zanichelli, Bologna, 2017, p. 302.

[16] Corsivo mio.

[17] Corsivo mio. Il *rischio sistemico* viene infatti definito come «un rischio specifico per le capacità di impatto elevato dei modelli di IA per finalità generali, avente un impatto significativo sul mercato dell'Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso, che può propagarsi su larga scala lungo l'intera catena del valore».

Questa soluzione, improntata alla flessibilità, va oltre il tradizionale (e rigido) principio di precauzione[18] alla cui applicazione consegue la statuizione di un divieto. La tecnica normativa basata sul principio di precauzione si differenzia dunque in maniera netta dall'approccio basato sul rischio, ma ne rappresenta – come si proverà a sostenere (*infra* par. 4.5 in relazione al rischio inaccettabile) – un necessario fattore complementare di organizzazione sistematica.

Nell'elaborazione dottrinale i due principi sono stati considerati alternativi ma non incompatibili[19] e l'opportunità/possibilità di ricorrere all'uno o all'altro criterio è stata anche legata al livello di conoscenza della specifica strumentazione regolamentata. È stato perciò evidenziato che la necessità di ricorrere all'approccio precauzionale per regolare l'utilizzo di una determinata tecnologia, sancendone un divieto o una restrizione, sia legata all'impossibilità di calcolare, con precisione e in anticipo, i potenziali gravi rischi legati al suo utilizzo[20]. Verrebbe così proibito non solo ciò che reca certamente danno alla salute o all'ambiente, ma anche ciò che non si è certi lo recherà[21].

L'approccio basato sul rischio, invece, verrebbe applicato nelle situazioni in cui si ritiene che i rischi possano essere chiaramente identificati e analizzati. Il processo valutativo potrebbe, nello specifico, essere scomposto nelle seguenti fasi: 1) identificazione dei rischi; 2) analisi del potenziale impatto degli stessi; 3) individuazione e adozione delle misure idonee a prevenire o limitare l'impatto; 4) successivo monitoraggio ed eventuale revisione. In tal modo si esercita una valutazione probabilistica volta a contemperare le esigenze di promozione dell'innovazione e la mitigazione dei rischi. In altri termini, si utilizza il rischio come strumento per stabilire le priorità e calibrare l'applicazione della legge in relazione al pericolo "reale"[22].

3. L'approccio basato sul rischio nel GDPR

In materia di tutela dei dati personali, a seguito della diffusione delle banche dati elettroniche, il concetto di rischio è stato preso in considerazione come elemento strutturale delle normative.

[18] Sul quale v. A. ROTA, *Sicurezza*, in M. NOVELLA, P. TULLINI (a cura di), *Lavoro digitale*, Giappichelli, Torino, 2022, p. 88 ss., con particolare riferimento al rischio da ignoto tecnologico. Sulle nozioni di prevenzione e precauzione cfr. S. BUOSO, *Principio di prevenzione e sicurezza sul lavoro*, Giappichelli, Torino, 2020, p. 7 ss.

[19] M. PERUZZI, *Intelligenza artificiale e lavoro*, *cit.*, p. 37. Sulla stessa scia: Council of Europe, Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108), 2017, Section IV, parr. 1 e 2; Commissione europea, Comunicazione sul principio di precauzione, 2 febbraio 2000 COM(2000)1. Esclude un'alternativa rigorosa anche A. MANTELERO, *Beyond Data. Human Rights, Ethical and Social Impact Assessment in AI*, Springer, Berlin, 2022, p. 50.

[20] M. BARBERA, *"La nave deve navigare"*, *cit.*, p. 12.

[21] *Ibidem*.

[22] Cfr. G. DE GREGORIO, P. DUNN, *The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age*, in *Common Market Law Review*, n. 59(2), 2022, p. 3.

Ciò inizia a intravedersi già nella direttiva 95/46/CE[23], ma si afferma emblematicamente nel GDPR, fondato sulla gestione graduata del rischio a seconda della sua gravità. Sin dal considerando 76, infatti, si evidenzia che «Il rischio dovrebbe essere considerato in base a una valutazione oggettiva mediante cui si stabilisce se i trattamenti di dati comportano un rischio o un rischio elevato».

Con il GDPR si consolida, pertanto, un sistema normativo non basato su una logica binaria (legale/illegale), ma su una logica scalare dell'analisi del rischio[24], che prevede la responsabilizzazione dei soggetti privati cui il regolamento delega il compito di individuare i mezzi adeguati a ottemperare ai requisiti legali.

Nell'articolato del GDPR l'approccio basato sul rischio si concreta in diverse previsioni, segnatamente: a) l'art. 24 «Responsabilità del titolare del trattamento», che – superando la previgente disciplina che forniva puntuali prescrizioni al riguardo[25] e ricorrendo al principio di *accountability* – lascia libero il titolare del trattamento di adottare discrezionalmente le misure più adeguate[26] anche in relazione ai «rischi aventi probabilità e gravità diverse» (par. 1); b) l'art. 25 «Protezione dei dati fin dalla progettazione e protezione dei dati per impostazione predefinita», che mette in relazione le soluzioni progettuali volte a garantire la tutela dei dati personali con il rischio dello specifico trattamento; c) l'art. 32 «Sicurezza del trattamento», per il quale le misure tecniche e organizzative devono garantire[27] un livello di sicurezza «adeguato al rischio»; d) gli artt. 35 «Valutazione d'impatto sulla protezione dei dati» e 36 «Consultazione preventiva», dai quali si desume un elemento ulteriore rispetto alla valutazione generalizzata del rischio, ossia la valutazione formalizzata e procedimentalizzata in caso di rischio elevato[28].

In ciascuna delle citate disposizioni non manca il riferimento alla protezione dei diritti fondamentali, elemento che il GDPR, nel primo articolo, individua come propria finalità. Emerge così come, nel contesto del trattamento dei dati personali, la regolamentazione miri a proteggere i diritti e le libertà

[23] Cfr. A. MANTELERO, *Il nuovo approccio*, *cit.*, p. 287 ss.

[24] V. R. GELLERT, *The Risk-Based Approach to Data Protection*, Oxford University Press, Oxford, 2020, p. 2: «The risk-based approach displays a very different logic. Namely, the granular, scalable, logic of risk analysis: it is less a matter of whether the processing is too risky or not, than a matter of determining how much risk one can take».

[25] In particolare l'allegato B del d.lgs. 30 giugno 2003, n. 196, espressamente abrogato dall'art. 27, co. 1, lett. *d*, del d.lgs. 10 agosto 2018, n. 101.

[26] Cfr. F. PIZZETTI, L. GRECO, *Art. 24*, in R. D'ORAZIO, G. FINOCCHIARO, O. POLLICINO, G. RESTA (a cura di), *Codice della privacy e data protection*, Giuffré, Milano, 2021, p. 405, i quali evidenziano che «il principio di accountability risponde … all'esigenza di fornire un parametro unico per tutte le situazioni di trattamento di dati personali ma, al tempo stesso, modulabile a seconda delle caratteristiche, dei compiti e delle finalità di ciascun titolare e dei trattamenti che pone in essere».

[27] Obbligo posto a carico del titolare e del responsabile del trattamento.

[28] Cfr. A. MANTELERO, *Art. 35*, in R. D'ORAZIO, G. FINOCCHIARO, O. POLLICINO, G. RESTA (a cura di), *Codice della privacy*, *cit.*, p. 542. Sul ruolo giocato dalle valutazioni d'impatto nella gestione dei rischi per i diritti dei lavoratori soggetti a *management* algoritmico nel GDPR, nell'*AI Act* e nella c.d. direttiva piattaforme – ossia la direttiva (UE) del 23 ottobre 2024, n. 2831 – v. G. GAUDIO, V*alutazioni d'impatto e management algoritmico*, in *Rivista giuridica del lavoro e della previdenza sociale*, n. 4, 2024, I, p. 538 ss.

fondamentali (cons. 51[29]; art. 24, par. 1[30]; art. 25, par. 1[31]; art. 35[32]), in particolare i dati che rivelano – per quanto di maggiore interesse in questa sede – l'appartenenza sindacale o, in caso di valutazione di aspetti personali, l'analisi o la previsione di elementi riguardanti il rendimento professionale (cons. 71)[33].

La prospettiva seguita è dal basso verso l'alto (*bottom-up*), nel senso che la valutazione del rischio e la scelta delle misure di mitigazione sono lasciate alla discrezione dei destinatari della regolamentazione, che sono tenuti a trovare l'equilibrio "ottimale" tra i contrastanti interessi. Ciò emerge dalle norme che disciplinano le responsabilità dei titolari del trattamento e dai principi di *privacy by design* e *by default* [34], i quali richiedono che i sistemi siano, rispettivamente, progettati in modo tale da prevenire il verificarsi dei rischi e trattare solo i dati personali necessari.

In altre parole, il principio di *accountability* (art. 5, par. 2) rappresenta la "chiave di volta" del regolamento europeo nella misura in cui impone al titolare del trattamento di essere in grado di dimostrare che siano state adottate tutte le misure tecniche e organizzative idonee a garantire la conformità del trattamento ai principi di liceità, trasparenza e correttezza[35].

Ad ogni buon conto, come si accennava in premessa, nel GDPR l'analisi del rischio risulta fondata sui diritti e sulla rilevanza degli stessi secondo la concezione "*right-based*": la tutela apprestata dal legislatore è strutturata come "protezione contro la violazione del diritto" e il concetto di rischio di violazione dei diritti viene distinto dal rischio di danno, menzionato – ad es. nel considerando 75 – solo come un'eventuale conseguenza legata al trattamento dei dati.

---

[29] «Meritano una specifica protezione i dati personali che, per loro natura, sono particolarmente sensibili sotto il profilo dei diritti e delle libertà fondamentali, dal momento che il contesto del loro trattamento potrebbe creare *rischi significativi per i diritti e le libertà fondamentali*» (corsivo mio).

[30] «Tenuto conto della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, nonché dei *rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche*, il titolare del trattamento mette in atto misure tecniche e organizzative adeguate per garantire, ed essere in grado di dimostrare, che il trattamento è effettuato conformemente al presente regolamento. Dette misure sono riesaminate e aggiornate qualora necessario» (corsivo mio).

[31] «Tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, come anche dei *rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche* costituiti dal trattamento, sia al momento di determinare i mezzi del trattamento sia all'atto del trattamento stesso il titolare del trattamento mette in atto misure tecniche e organizzative adeguate, quali la pseudonimizzazione, volte ad attuare in modo efficace i principi di protezione dei dati, quali la minimizzazione, e a integrare nel trattamento le necessarie garanzie al fine di soddisfare i requisiti del presente regolamento e tutelare i diritti degli interessati» (corsivo mio).

[32] «Quando un tipo di trattamento, allorché prevede in particolare l'uso di nuove tecnologie, considerati la natura, l'oggetto, il contesto e le finalità del trattamento, *può presentare un rischio elevato per i diritti e le libertà delle persone fisiche*, il titolare del trattamento effettua, prima di procedere al trattamento, una valutazione dell'impatto dei trattamenti previsti sulla protezione dei dati personali. Una singola valutazione può esaminare un insieme di trattamenti simili che presentano rischi elevati analoghi» (corsivo mio).

[33] Sul tema v. L. D'ARCANGELO, *La tutela del lavoratore nel trattamento dei dati personali*, Aracne, Roma, 2024.

[34] Di cui all'art. 25 GDPR. Sul punto cfr. V. BRINO, *Privacy by default (voce)* e *Privacy by design (voce)* in S. BORELLI, V. BRINO, C. FALERI, L. LAZZERONI, L. TEBANO, L. ZAPPALÀ (a cura di), *Lavoro e tecnologie. Dizionario del diritto del lavoro che cambia*, Giappichelli, Torino, 2022, risp. p. 173 ss. e p. 170 ss.

[35] Così L. TEBANO, *Poteri datoriali e dati biometrici nel contesto dell'AI Act*, in q. *Rivista*, n. 25, 2023, pp. 202-203.

4. L'approccio basato sul rischio nell'*AI Act*

La regolazione europea in materia di intelligenza artificiale individua nell'approccio basato sul rischio una precisa scelta di metodo[36], nonostante possa sembrare una contraddizione l'idea di cristallizzare – seppur nei termini in cui si dirà a breve[37] – le soglie di rischio (cioè di identificarle *ex ante* in maniera chiara) per regolamentare una tecnologia che si distingue proprio per l'opaco funzionamento e per la capacità di modificarsi e migliorarsi automaticamente attraverso l'esperienza e l'utilizzo dei dati; caratteristiche ribadite dagli orientamenti della Commissione – approvati il 6 febbraio 2025 – sulla definizione di sistema di IA[38].

Il regolamento europeo sull'IA, fornendo un livello di protezione minimo e orizzontale[39], attraverso l'approccio *risk-based* [40] si prefigge lo scopo di migliorare il funzionamento del mercato interno e

[36] Sin dal Libro bianco sull'IA del febbraio 2020 – Commissione europea, 19.2.2020 COM(2020) 65 *final*, Libro Bianco sull'intelligenza artificiale - Un approccio europeo all'eccellenza e alla fiducia – la Commissione europea ha chiarito che quest'approccio risulta opportuno «per garantire la proporzionalità dell'intervento normativo» (par. 4, sez. C). La necessità di ricorrervi è stata ribadita anche nella Risoluzione del Parlamento europeo del 20 ottobre 2020 – recante raccomandazioni alla Commissione concernenti il quadro relativo agli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie correlate (2020/2012(INL)), punto *S* introduzione e paragrafi 12-16 –, che, al contempo, auspica un intervento «in linea con il principio di precauzione che guida la legislazione dell'UE» (in una logica che confermerebbe la complementarità di cui *supra* al par. 2 tra principio di precauzione e approccio basato sul rischio). E si ritrova pure nella Dichiarazione europea sui diritti e i principi digitali per il decennio digitale (2023/C 23/01) proclamata da Parlamento europeo, Consiglio e Commissione, nella quale le Istituzioni europee si impegnano a «garantire che l'utilizzo dell'intelligenza artificiale sul luogo di lavoro sia trasparente e segua un approccio basato sul rischio e che siano adottate le corrispondenti misure di prevenzione per mantenere un ambiente di lavoro sicuro e sano».

[37] E in particolare con riferimento al giudizio di significatività del rischio (art. 6) e al potere della Commissione di modificare l'elenco di sistemi ad alto rischio di cui all'allegato III (art. 7).

[38] Cfr. European Commission, *Annex to the Communication to the Commission. Approval of the content of the draft Communication from the Commission - Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)*, C(2025) 924 final, 6.2.2025; orientamenti che, in forza della previsione di cui all'art. 96, par. 1, lett. *f*, riguardano l'applicazione dell'art. 3, pt. 1, il quale definisce il sistema di IA come «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali». Sulla definizione di sistema di IA v. A. ODDENINO, *Intelligenza artificiale e tutela dei diritti fondamentali: alcune notazioni critiche sulla recente Proposta di Regolamento della UE, con particolare riferimento all'approccio basato sul rischio e al pericolo di discriminazione algoritmica*, in A. PAJNO, F. DONATI, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione? Diritti fondamentali, dati personali e regolazione*, vol. I, Il Mulino, Bologna, 2022, p. 183 ss. Sulla difficoltà di determinare *a priori* il rischio con specifico riferimento al rischio di danno v. C. CRISTOFOLINI, *Navigating the impact of AI systems in the workplace: strengths and loopholes of the EU AI Act from a labour perspective*, in *Italian Labour Law e-Journal*, n. 1, 2024, p. 84, ove si precisa che: «it can be challenging to properly classify and determine the "significance of the risk of harm" in the development phase».

[39] A. ADINOLFI, *L'intelligenza artificiale tra rischi di violazione dei diritti fondamentali e sostegno della loro promozione: considerazioni sulla (difficile) costruzione di un quadro normativo dell'Unione*, in A. PAJNO, F. DONATI, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto*, *cit.*, p. 127.

Il regolamento, ai sensi dell'art. 2, par. 11, infatti, «non osta a che l'Unione o gli Stati membri mantengano o introducano disposizioni legislative, regolamentari o amministrative più favorevoli ai lavoratori in termini di tutela dei loro diritti in relazione all'uso dei sistemi di IA da parte dei datori di lavoro, o incoraggino o consentano l'applicazione di contratti collettivi più favorevoli ai lavoratori».

[40] S. CIUCCIOVINO, *Risorse umane e intelligenza artificiale alla luce del regolamento (UE) 2024/1689, tra norme legali, etica e codici di condotta*, in *Diritto delle Relazioni Industriali*, n. 3, 2024, p. 587, precisa che «il regolamento opta per una tecnica in parte immediatamente prescrittiva ed in parte basata sulla responsabilizzazione (*accountability*) e sulla valutazione

promuovere la diffusione di un'IA antropocentrica e affidabile, garantendo un livello elevato di protezione della salute, della sicurezza e dei diritti sanciti dalla Carta dei diritti fondamentali[41] (art. 1). La configurazione di un'IA improntata all'antropocentrismo viene dunque perseguita non solo attraverso le regole definite dal regolamento, ma anche tramite i «principi etici non vincolanti» elaborati dall'*AI HLEG* (*High-Level Expert Group on Artificial Intelligence*)[42] richiamati nel cons. 27[43] e per mezzo delle disposizioni di diritto dell'Unione alle quali l'*AI Act* rinvia, che restano impregiudicate e richiedono un'operazione di coordinamento.

L'approccio basato sul rischio si concretizza nella possibilità di adattare la tipologia e il contenuto delle regole all'intensità e alla portata dei rischi che possono essere generati dai sistemi di IA, vietando determinate pratiche di IA inaccettabili (capo II, art. 5 ss.) e stabilendo requisiti di liceità per i sistemi di IA ad alto rischio (capo III, art. 6 ss.) e obblighi per gli operatori interessati, nonché obblighi di trasparenza per determinati sistemi di IA[44].

Muovendo dal considerando 46 dell'*AI Act* – il quale sancisce che è opportuno limitare i sistemi di IA identificati come ad alto rischio a quelli che hanno un impatto nocivo significativo sulla salute, sulla sicurezza e sui diritti fondamentali delle persone nell'Unione (con puntuale riferimento ai diritti riconosciuti dalla Carta nel considerando 48[45]) – risulta possibile ricostruire i vari livelli di rischio individuati dal regolamento attraverso il prisma dei diritti fondamentali. In questo modo, secondo la logica "a livelli" che caratterizza l'approccio *risk-based*, si possono identificare quattro macrocategorie di

preventiva del rischio (*risk-based*) dei soggetti obbligati», permettendo così la massima adattabilità delle regole alle specifiche applicazioni.

[41] In relazione all'applicazione della Carta di Nizza all'*AI Act* (*rectius* ai soggetti menzionati nel regolamento), v. M. DELFINO, *Fundamental rights in the age of Robotics, cit.,* p. 124 ss.
In generale, sul ruolo dei diritti fondamentali nell'*AI Act*, v. E. CIRONE, *L'AI Act e l'obiettivo (mancato?) di promuovere uno standard globale per la tutela dei diritti fondamentali*, in *Quaderni AISDUE*, fascicolo speciale n. 2, 2024, p. 12 ss., la quale li considera «scheletro dell'intero impianto normativo».

[42] Cfr. S. CIUCCIOVINO, *Risorse umane e intelligenza artificiale*, cit., p. 591, che evidenzia la necessità di osservarli «almeno come canoni interpretativi delle norme europee e del regolamento pertinenti in materia di IA».

[43] Il quale prevede che: «Sebbene l'approccio basato sul rischio costituisca la base per un insieme proporzionato ed efficace di regole vincolanti, è importante ricordare gli orientamenti etici per un'IA affidabile del 2019 elaborati dall'AI HLEG indipendente nominato dalla Commissione. In tali orientamenti l'AI HLEG ha elaborato sette principi etici non vincolanti per l'IA che sono intesi a contribuire a garantire che l'IA sia affidabile ed eticamente valida. I sette principi comprendono: intervento e sorveglianza umani, robustezza tecnica e sicurezza, vita privata e governance dei dati, trasparenza, diversità, non discriminazione ed equità, benessere sociale e ambientale e responsabilità. Fatti salvi i requisiti giuridicamente vincolanti del presente regolamento e di qualsiasi altra disposizione di diritto dell'Unione applicabile, tali orientamenti contribuiscono all'elaborazione di un'IA coerente, affidabile e antropocentrica, in linea con la Carta e con i valori su cui si fonda l'Unione». Sull'approccio umano-centrico all'intelligenza artificiale, v. T. TREU, *Intelligenza Artificiale (IA): integrazione o sostituzione del lavoro umano?*, in *WP CSDLE "Massimo D'Antona".IT,* 487/2024, p. 15 ss.

[44] Così si legge nel considerando 26.

[45] Il quale, tra gli altri, fa riferimento al diritto alla dignità umana, al rispetto della vita privata e della vita familiare alla protezione dei dati personali, alla libertà di espressione e di informazione, alla libertà di riunione e di associazione e al diritto alla non discriminazione, ai diritti dei lavoratori, ai diritti delle persone con disabilità e all'uguaglianza di genere.

sistemi di IA, a seconda del potenziale impatto sui diritti fondamentali[46]. Si tratta di una quadripartizione che non trova puntuale riscontro nell'*AI Act*, ma che si rinviene anche nei comunicati stampa della Commissione europea[47] in quanto funzionale in chiave descrittiva.

Anzitutto, in un primo gruppo, cui corrisponde un *rischio minimo*, trovano residuale collocazione i sistemi il cui utilizzo non sia considerato impattante sui diritti fondamentali e, per questo, non sia necessario prevedere specifiche forme di tutela (es. videogiochi abilitati all'intelligenza artificiale o filtri *antispam*); potendosi al più rinviare alla spontanea adozione dei codici di condotta di cui all'art. 95[48].

In un secondo gruppo, cui corrisponde un *rischio limitato*, rientrano i sistemi di intelligenza artificiale per i quali il regolamento considera sufficiente intervenire sul piano della trasparenza (ad es. l'informativa sull'utilizzo del sistema per le *chatbot*) per scongiurare l'impatto negativo sull'area dei diritti fondamentali. Si tratta di sistemi che, pur non rientrando tra quelli ad alto rischio (di cui al gruppo successivo), sono soggetti alle prescrizioni "trasversali" di cui al capo IV del regolamento (art. 50), rubricato «Obblighi di trasparenza per i fornitori e i deployer di determinati sistemi di IA», il quale si rivolge a specifiche ipotesi come i sistemi di IA destinati a interagire direttamente con le persone fisiche (art. 50, par. 1).

Il terzo (*infra* 4.1) e il quarto gruppo (*infra* 4.5), corrispondenti rispettivamente all'individuazione di un rischio alto o inaccettabile, rappresentano il cuore della regolamentazione e, per tale ragione, richiedono un'analisi circostanziata.

4.1. Sistemi di IA ad alto rischio e *autoreferenzialità classificatoria*

In un terzo gruppo rientrano i sistemi contemplati al capo III (art. 6 ss.), al cui impiego corrisponde un *alto rischio* di pregiudicare i diritti fondamentali. Il regolamento non definisce l'alto rischio ma si limita a considerare ad alto rischio i sistemi di IA riconducibili alla legislazione armonizzata di cui all'allegato I in materia di sicurezza dei prodotti (art. 6, par. 1)[49] e ai settori e ai casi d'uso indicati nell'elenco di cui allegato

[46] Individua, invece, nell'ipotesi dei sistemi di IA ad alto rischio, il ricorso tanto alla logica della prevenzione quanto della mitigazione, I. SENATORI, in S. SCAGLIARINI, I. SENATORI (a cura di), *Lavoro, Impresa e Nuove Tecnologie dopo l'AI Act*, *Quaderno Fondazione Marco Biagi*, n. 17, 2024, p. 7: «La logica seguita per la regolazione dei sistemi ad alto rischio è quella della prevenzione (dove il rischio è quello della violazione di un diritto fondamentale), o, qualora il rischio sia ineliminabile, della sua mitigazione. La tecnica è invece quella della imposizione di una serie di obblighi procedurali, il cui corretto assolvimento dovrebbe liberare da responsabilità nel caso in cui si verifichi un evento pregiudizievole».

[47] La quadripartizione, a quanto consta, è stata elaborata dalla stessa Commissione, che continua ad adottarla anche in seguito dell'approvazione dell'AI Act, (ultimo aggiornamento 18 Febbraio 2025). Tale ricostruzione, invero, semplifica e, al contempo, va oltre il dato normativo che dei quattro livelli di rischio espressamente contempla solo l'*alto rischio* e (ma solo nel considerando 179) il *rischio inaccettabile* (mentre nell'articolato si fa riferimento alle pratiche vietate). Questa classificazione è stata ripresa da diversi autori, per tutti, M. EBERS, *Truly Risk-Based Regulation of Artificial Intelligence - How to Implement the EU's AI Act*, *EU Law Working Papers Transatlantic Technology Law Forum* n. 101, 2024. p. 4.

[48] S. CIUCCIOVINO, *Risorse umane e intelligenza artificiale*, *cit.,* p. 602 ss.

[49] V. art. 6, par. 1, il quale richiede il soddisfacimento di due condizioni: il sistema è destinato a essere utilizzato come componente di sicurezza di un prodotto o è esso stesso un prodotto disciplinato dalla normativa di armonizzazione dell'Unione elencata nell'allegato I – ad esempio regolamento (UE) 2016/425 del Parlamento europeo e del Consiglio, del 9 marzo 2016, sui dispositivi di protezione individuale – (lett. *a*) ed è soggetto a una valutazione della conformità da

III (al quale rinvia l'art. 6, par. 2). Tra questi richiama qui l'attenzione il settore «Occupazione, gestione dei lavoratori e accesso al lavoro autonomo» (pt. 4) per quanto concerne i sistemi di IA destinati a essere utilizzati per l'assunzione o la selezione di persone fisiche e gli strumenti di gestione algoritmica del rapporto di lavoro con funzione decisionale, allocativa o valutativa. Alla Commissione – ai sensi dell'art. 7 – è conferito il potere di modificare l'allegato III, aggiungendo o modificando i casi d'uso, qualora determinati sistemi di IA dovessero presentare «un rischio di danno per la salute e la sicurezza, o di impatto negativo sui diritti fondamentali … equivalente o superiore al rischio di danno o di impatto negativo presentato dai sistemi di IA ad alto rischio di cui all'allegato III».

Sin qui l'impostazione regolamentare potrebbe apparire graniticamente *top-down*: come ampiamente sostenuto in dottrina prima della definitiva approvazione del regolamento, la valutazione dei livelli di rischio sembrerebbe il frutto di una scelta operata a monte dal legislatore. Tuttavia, la riconducibilità all'elencazione di cui all'allegato III non rappresenta garanzia di applicazione del corrispondente regime giuridico[50]. Infatti, in forza della deroga di cui all'art. 6, par. 3, «un sistema di IA di cui all'allegato III non è considerato ad alto rischio se non presenta un *rischio significativo di danno* per la salute, la sicurezza o i diritti fondamentali delle persone fisiche»[51] in una serie di condizioni tra le quali spicca – per la genericità della formulazione – l'essere destinato a eseguire un compito procedurale limitato[52]. L'aggettivo "significativo" – sembra evidente – è riferito alla valutazione della probabilità del verificarsi del danno e non alla sua gravità: è come se il legislatore avesse detto, in maniera tautologica, che il sistema è configurabile come ad alto rischio se il rischio è alto. In forza di questa deroga, ad esempio, un sistema decisionale automatizzato utilizzato per l'assegnazione dei compiti ai lavoratori (quindi considerato ad alto rischio ai sensi dell'art. 6, par. 2, perché riconducibile all'elenco di cui all'allegato III) potrebbe essere qualificato dal fornitore come a rischio limitato dinanzi non solo alla limitatezza del compito, ma anche (e soprattutto) all'asserita assenza di *rischio significativo di danno* per i diritti fondamentali della persona.

parte di terzi ai fini dell'immissione sul mercato o della messa in servizio ai sensi della medesima normativa di armonizzazione (lett. *b*).

[50] Anche nelle citate *Guidelines* del 6 febbraio 2025 (v. *supra* nota 38) sulla definizione di sistema di IA si precisa che: «(63) *Only certain AI systems are subject to regulatory obligations and oversight under the AI Act. The AI Act's risk-based approach means that only those systems giving rise to the most significant risks to fundamental rights and freedoms will be subject to its prohibitions laid down in Article 5 AI Act, its regulatory regime for high-risk AI systems covered by Article 6 AI Act and its transparency requirements for a limited number of pre-defined AI systems laid down in Article 50 AI Act. The vast majority of systems, even if they qualify as AI systems within the meaning of Article 3(1) AI Act, will not be subject to any regulatory requirements under the AI Act*».

[51] Corsivo mio. La disposizione aggiunge: «anche nel senso di non influenzare materialmente il risultato del processo decisionale».

[52] D'altronde si potrebbe arrivare a sostenere che ogni compito definito sia di per sé limitato. Altre ipotesi sono il caso in cui il sistema sia destinato a «migliorare il risultato di un'attività umana precedentemente completata»; a «rilevare schemi decisionali o deviazioni da schemi decisionali precedenti e non (sia) finalizzato a sostituire o influenzare la valutazione umana precedentemente completata senza un'adeguata revisione umana» o a «eseguire un compito preparatorio per una valutazione pertinente ai fini dei casi d'uso elencati nell'allegato III».

L'ambiguità di questa disposizione investe l'intera architettura normativa dell'*AI Act* e nega la qualificazione dell'impostazione regolamentare come *top-down,* anche se qualche aspettativa di chiarimento può essere riposta negli orientamenti della Commissione – attesi entro il 2 febbraio 2026 – che daranno indicazioni per l'attuazione pratica delle Regole di classificazione.

Ciò vuol dire che la qualificazione del sistema, di fatto, non è data dalla corrispondenza ai predefiniti (per quanto modificabili) casi d'uso di cui all'allegato III, ma è rimessa – in ultima istanza – al fornitore[53] che si vede investito di un'ampia discrezionalità.

Emerge così il tema – che in seguito sarà affrontato con accezioni diverse – della "autoreferenzialità": allo stesso produttore viene affidato – come descritto – il potere di qualificare o meno il sistema come ad alto rischio, con tutto ciò che ne consegue in termini di disciplina da applicare.

Quest'*autoreferenzialità classificatoria* potrebbe però essere ridimensionata nell'ipotesi di utilizzo di sistemi di IA che facciano ricorso alla profilazione. Sin dal considerando 53 si sottolinea, infatti, che «In ogni caso, è opportuno ritenere che i sistemi di IA utilizzati (per gli) usi ad alto rischio elencati nell'allegato del presente regolamento comportino rischi significativi di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche se il sistema di IA implica la profilazione ai sensi dell'articolo 4, punto 4, del regolamento (UE) 2016/679». Nell'articolato, poi, l'art. 6, par. 3, ultimo capoverso, sancisce che «un sistema di IA di cui all'allegato III è sempre considerato ad alto rischio qualora esso effettui profilazione di persone fisiche», vale a dire – per il richiamo al GDPR nell'art. 3, pt. 52[54] – qualsiasi forma di trattamento automatizzato di dati personali finalizzata alla valutazione di determinati aspetti personali relativi a una persona fisica, tra cui l'analisi e la previsione di aspetti riguardanti il rendimento professionale, ma anche l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona[55].

Questa disposizione potrebbe limitare significativamente l'ambito applicativo della deroga rappresentando un importante presidio normativo per garantire che la maggior parte dei sistemi impiegati in ambito lavorativo sia classificata come ad alto rischio, siccome è molto probabile che la profilazione sia connaturata al funzionamento di gran parte dei sistemi utilizzati nei contesti lavorativi[56].

---

[53] Soggetto che – *ex* art. 3, pt. 3 – sviluppa o fa sviluppare il sistema (o il modello di IA per finalità generali) e lo immette sul mercato o lo mette in servizio.

[54] Che definisce la profilazione come «la profilazione quale definita all'articolo 4, punto 4), del regolamento (UE) 2016/679».

[55] Cfr. art. 4, pt. 4, del GDPR.

[56] Secondo M. PERUZZI, *Intelligenza artificiale e lavoro: l'impatto dell'AI Act nella ricostruzione del sistema regolativo UE di tutela*, in M. BIASI (a cura di), *Diritto del lavoro e intelligenza artificiale*, Giuffrè, Milano, 2024, p. 128, è molto probabile che la profilazione si verifichi in caso di utilizzo di processi automatizzati di monitoraggio o decisionali sul lavoro e, pertanto, è «tendenzialmente da escludere una classificazione come "a basso rischio" dei sistemi destinati a essere utilizzati nell'esercizio di poteri datoriali e nei confronti dei lavoratori»; A parere invece di A. ALAIMO, *Il Regolamento sull'Intelligenza Artificiale. Un treno al traguardo con alcuni vagoni rimasti fermi*, in q. *Rivista*, n. 25, 2024, p. 235, questa previsione permette la definizione come ad alto rischio della maggior parte dei sistemi utilizzabili nella fase che precede l'instaurazione dei rapporti di lavoro, «Viceversa, i sistemi impiegabili in corso di rapporto potrebbero beneficiare di

Al riguardo viene in rilievo una questione di taglio strettamente esegetico: l'automatismo profilazione-alto rischio è legato alla proposizione subordinata «Fatto salvo il primo comma»[57] (art. 6, par. 3, *incipit* dell'ultimo capoverso), ove il riferimento al primo comma non può che intendersi al primo comma del paragrafo, ossia alla deroga per assenza di rischio significativo di danno[58]. Di primo acchito si potrebbe pensare che l'automatismo in parola non operi in presenza della deroga "fatta salva". Così opinando, evidentemente, si svuoterebbe di senso tanto il citato considerando quanto la stessa disposizione[59]. Diversamente, se si legge il testo redatto in lingua inglese del regolamento, la subordinata si rivela foriera di una contro-deroga. L'*incipit* in inglese dell'ultimo capoverso è «*Notwithstanding the first subparagraph*», ossia "nonostante il" o "in deroga al" primo capoverso (del paragrafo in questione), il quale prevede la deroga (stante l'assenza di rischio significativo di danno) alla regola generale (di cui al par. 2) che qualifica ad alto rischio i sistemi riconducibili all'elencazione dell'allegato III. In questi termini, la profilazione escluderebbe ogni sorta di valutazione discrezionale del fornitore e tutti i sistemi riconducibili all'allegato III dovrebbero essere qualificati come ad alto rischio. È dunque da preferire la lettura del testo in lingua inglese se – ripeto – non si vuole svuotare di senso la disposizione[60]; tuttavia, anche nell'attesa di un eventuale *corrigendum*, risulta consigliabile avanzare con circospezione.

4.2. (Segue) Conformità statica e *autoreferenzialità certificatoria*

Oltrepassando la – più o meno ampia – *autoreferenzialità classificatoria,* l'autoreferenzialità (qui di nuovo) del fornitore assume una differente sfaccettatura in relazione alla procedura di certificazione della conformità per i sistemi di IA ad alto rischio fissati dalla sez. 2 del capo III dell'*AI Act* (art. 8 e ss.).

Al riguardo pare opportuno distinguere tra conformità "statica", che riguarda l'immissione del prodotto/sistema di IA ad alto rischio nel mercato, e conformità "dinamica", che invece concerne gli sviluppi applicativi e coinvolge la figura del *deployer* (*infra* par. 4.3).

In relazione al primo profilo – semplificando, certificazione per la commercializzazione – si ricorda che la sez. 2 del capo III prevede: istituzione, attuazione, documentazione e mantenimento di un sistema di

valutazioni di non significatività del rischio (capaci di sottrarli, quindi, ai relativi obblighi)», sempre che sia possibile dimostrare le condizioni richiamate dalla medesima disposizione, fra le quali spicca la sorveglianza umana. Sui presidi informativi e sul ruolo del *deployer* cfr. L. TEBANO*, Intelligenza Artificiale e datore di lavoro: scenari e regole*, in *Diritti Lavori Mercati*, n. 3, 2024, p. 453 ss.

[57] Testualmente: «Fatto salvo il primo comma, un sistema di IA di cui all'allegato III è sempre considerato ad alto rischio qualora esso effettui profilazione di persone fisiche».

[58] L'ultimo capoverso dell'art. 6, par. 3, (riportato nella nota precedente) non può che essere riferito al primo comma del terzo paragrafo dato che non avrebbe senso il riferimento al primo comma dell'articolo (ossia al par. 1) che considera ad alto rischio i sistemi di IA riconducibili alla legislazione armonizzata di cui all'allegato I in materia di sicurezza dei prodotti. Questa lettura è anche in linea con il citato considerando 53.

[59] Non vi sarebbe riscontro nell'articolato a quanto prospettato nel considerando. Inoltre, dando preminenza, in ogni caso, alla derogabilità per assenza di rischio significativo di danno si vanificherebbe la portata dell'art. 6, par. 3.

[60] Coerentemente con l'avverbio «sempre», presente nella formulazione della norma.

gestione dei rischi (art. 9); sottoposizione a pratiche di *governance* e gestione per i *set* di dati di addestramento, convalida e prova (art. 10), redazione della documentazione tecnica di conformità (art. 11), registrazione automatica degli eventi (*log*) durante il funzionamento del sistema (art. 12), trasparenza e fornitura di informazioni ai *deployer* (art. 13), sorveglianza umana (art.14), accuratezza, robustezza, e cibersicurezza (art. 15).

Tra le elencate disposizioni, anche in funzione dello sviluppo della presente analisi, merita specifica attenzione il sistema di gestione dei rischi di cui all'art. 9, il quale si caratterizza per una propensione verso la prospettiva dinamica, essendo un «processo iterativo continuo»[61] le cui misure di gestione «sono tali che i pertinenti rischi residui associati a ciascun pericolo nonché il rischio residuo complessivo dei sistemi di IA ad alto rischio sono considerati accettabili»[62]. Tuttavia, la portata della presente disposizione è limitata[63]. Come recita il par. 3, infatti: «I rischi di cui al presente articolo riguardano solo quelli che possono essere ragionevolmente attenuati o eliminati attraverso lo sviluppo o la progettazione del sistema di IA ad alto rischio o la fornitura di informazioni tecniche adeguate»; inoltre devono essere «rischi noti e ragionevolmente prevedibili»[64].

A ben vedere, il regolamento individua a grandi linee le caratteristiche del "prodotto intelligenza artificiale" senza definire parametri tecnici dettagliati e vincolanti: i requisiti appena elencati si configurano più come obiettivi o indicazioni generali che come risultati concreti e misurabili, lasciando ampi margini di discrezionalità alle soluzioni tecniche adottabili dai fornitori. Emblematico è l'art. 13, par. 1, il quale stabilisce che il funzionamento dei sistemi di IA ad alto rischio deve essere «sufficientemente trasparente da consentire ai deployer di interpretare l'output del sistema e utilizzarlo adeguatamente», non chiarendo – ad esempio – se il livello di sufficienza possa essere soddisfatto dall'indicazione dei parametri che concorrono alla definizione dell'*output*, la cui specifica incidenza risulta non definita[65].

---

[61] «Il sistema di gestione dei rischi è inteso come un processo iterativo continuo pianificato ed eseguito nel corso dell'intero ciclo di vita di un sistema di IA ad alto rischio, che richiede un riesame e un aggiornamento costanti e sistematici» (par. 2).

[62] Art. 9, par. 5. Al riguardo, L. ZAPPALÀ, *Sistemi di IA ad alto rischio e ruolo del sindacato alla prova del* risk-based approach, in *Labour & Law Issues*, n. 1, 2024, p. 58, precisa che «Si tratta di un approccio che, inevitabilmente, non tende a eliminare i rischi, ma a ridurli al minimo o al punto tale da considerarli "accettabili" nel bilanciamento con i diritti fondamentali e alla luce dell'applicazione del test di proporzionalità … Un approccio comunque che, tendenzialmente, pur non presupponendo esiti scontati, consente di introdurre a monte una valutazione sulla "tollerabilità" dei rischi imprevedibili sottesi all'utilizzo delle nuove tecnologie ad alta complessità e velocità di cambiamento, quindi nemmeno astrattamente tutti prevedibili per via legislativa».

[63] Limitazione «duplice» (nei termini in cui si dirà) evidenziata anche da R.D. COSTA, *Lavoro e gestione del rischio all'epoca dell'intelligenza artificiale. Note a margine dell'AI ACT*, in R. SANTUCCI, A. TROJSI (a cura di), *Diritto del lavoro e intelligenza artificiale tra rischi e benefici*, Giappichelli, Torino, 2025, p. 96.

[64] Art. 9, par. 2, lett. *a.*

[65] Sulle diverse accezioni del principio di trasparenza (di interazione, di funzionamento e di impatto) cfr. S. CIUCCIOVINO, *Risorse umane e intelligenza artificiale*, *cit.,* p. 596 ss.

Questa soluzione potrebbe apparire coerente con l'esigenza di non porre limitazioni allo sviluppo tecnologico, ma diviene problematica nel momento in cui allo stesso fornitore viene attribuito il potere di certificare la conformità agli standard regolamentari. La certificazione di conformità al regolamento dei sistemi di IA ad alto rischio scaturisce, per l'appunto, da una procedura di valutazione basata sul controllo interno (autovalutazione)[66] di cui all'allegato VI, che non prevede il coinvolgimento di un soggetto terzo (art. 43). Solo nel caso di sistemi ad alto rischio che utilizzano dati biometrici – là dove naturalmente non si ricada nelle pratiche vietate di cui al successivo gruppo – si prevede (anche se non è sempre necessario) il coinvolgimento di un organismo terzo per la valutazione della conformità (definito organismo notificato)[67]. Per evitare la coincidenza tra i soggetti che certificano/vigilano sul bilanciamento degli interessi e coloro che possono beneficiare di un eventuale squilibrio – screditando il ricorso all'autocertificazione – sarebbe stato auspicabile ampliare le ipotesi di coinvolgimento delle autorità di notifica in parola[68], ferma restando la previsione delle sanzioni amministrative pecuniarie (*ex* art. 99) cui il fornitore è soggetto in caso di erronea classificazione finalizzata all'elusione dei citati requisiti di cui al capo III, sez. 2 (art. 80, par. 7); soluzione da sola inadeguata perché subordina la tutela dei diritti fondamentali a una mera valutazione economica di costo/beneficio per le aziende senza una loro protezione sostanziale.

Da qui, l'emersione dell'*autoreferenzialità certificatoria* del fornitore. Un argine ad essa, tuttavia, potrebbe rinvenirsi nel complessivo disegno dell'*AI Act* attraverso l'adozione delle c.d. "norme armonizzate". Almeno così sembra suggerire l'art. 40, par. 1, prevedendo che i sistemi di IA ad alto rischio (oltre che i modelli di IA per finalità generali[69]) conformi alle norme armonizzate si presumono conformi ai requisiti

[66] Sempre al fornitore spetta dichiarare la conformità del sistema ad alto rischio, oltre che all'*AI Act*, alle pertinenti normative dell'UE (art. 47, rubricato «Dichiarazione di conformità UE») e registrarlo nella banca dati dell'UE (art. 71).

[67] Il regolamento prevede, infatti, l'istituzione di autorità nazionali deputate a vigilare sull'applicazione degli obblighi e a sanzionarne le violazioni, nello specifico: la designazione di un'autorità di notifica (ossia – ai sensi dell'art. 3, pt. 19 – l'autorità nazionale responsabile dell'istituzione e dell'esecuzione delle procedure necessarie per la valutazione, la designazione e la notifica degli organismi di valutazione della conformità e per il loro monitoraggio) e di almeno un'autorità di vigilanza del mercato. Al riguardo, il d.d.l. n. 1146 (Disposizioni e deleghe al Governo in materia di intelligenza artificiale), approvato in Senato il 20 marzo 2025, designa l'Agenzia per l'Italia digitale (AgID) come autorità di notifica (art. 20, co. 1, lett. *a*) e l'Agenzia per la cybersicurezza nazionale (ACN) come autorità di vigilanza (art. 20, co. 1, lett. *b*). Nel parere circostanziato (C(2024) 7814) del 5 novembre 2024 della Commissione europea sul d.d.l. in questione viene ricordato che le autorità nazionali competenti devono possedere lo stesso livello di indipendenza previsto dalla direttiva (UE) n. 2016/680 per le autorità preposte alla protezione dei dati nelle attività delle forze dell'ordine, nella gestione delle migrazioni e controllo delle frontiere, nell'amministrazione della giustizia e nei processi democratici. Il d.d.l. si distingue per la previsione, all'art. 12, di un «Osservatorio sull'adozione di sistemi di intelligenza artificiale nel mondo del lavoro».

[68] Su queste preoccupazioni, v. U. GARGIULO, *Intelligenza Artificiale e poteri datoriali: limiti normativi e ruolo dell'autonomia collettiva*, in q. *Rivista*, n. 29, 2023, p. 172 ss., il quale, in particolare, evidenzia che «a mancare sono dispositivi di controllo e intervento che possano fare da contraltare ai rischi e, eventualmente, agli abusi dell'intelligenza artificiale: contesti che potremmo definire di "contropotere" rispetto a un esercizio del potere che, esaltando l'asimmetria delle posizioni nel contratto, risulti agevolato dai sistemi di AI, a detrimento di diritti fondamentali della controparte» (p. 180).

[69] Sui quali si tornerà *infra* par. 5.

regolamentari[70], qualora i riferimenti delle suddette norme siano stati pubblicati nella Gazzetta ufficiale dell'Unione europea conformemente al regolamento (UE) n. 1025/2012[71]. L'adozione dei criteri tecnici specifici sarebbe in questo modo demandata alle Organizzazioni europee di standardizzazione (ESO) attraverso il recepimento degli "standard" da parte della Commissione europea, al cui sviluppo potrebbero contribuire anche le organizzazioni sindacali[72]; soluzione che mostra però una serie di problematicità, su tutte: la natura giuridica delle ESO (con i conseguenti potenziali conflitti d'interesse) e la vaghezza dei principi regolamentari di riferimento[73].

4.3. (Segue) Conformità dinamica: l'obbligo datoriale di monitorare il sistema

Il regolamento prevede però una valutazione eventuale – che prescinde dalla qualificazione del sistema di IA e che ben si appresta a intervenire nella fase successiva alla messa in commercio – da parte dell'autorità di vigilanza del mercato dei singoli Stati membri. Ai sensi dell'art. 79, par. 2, infatti, l'autorità in parola «effettua una valutazione del sistema di IA interessato per quanto riguarda la sua conformità a tutti i requisiti e gli obblighi» dell'*AI Act* ogni qual volta «abbia un motivo sufficiente per ritenere che un sistema di IA presenti un rischio»[74] per la salute o la sicurezza o per i diritti fondamentali delle persone e, qualora dovesse riscontrare la non conformità, è tenuta a chiedere «al pertinente operatore di adottare tutte le misure correttive adeguate … (di) ritirarlo dal mercato o (di) richiamarlo». A tal proposito si sottolinea che la norma utilizza il termine operatore, figura che – ai sensi dell'art. 3, pt. 8 – può individuare tanto il fornitore quanto il *deployer*75, ossia – per quanto di maggiore interesse in questa sede – il datore di lavoro, che potrebbe essere il destinatario dei provvedimenti dell'autorità di vigilanza.

Significativamente, lo stesso *deployer* – ai sensi dell'art. 26, par. 5[76] – è tenuto a informare l'autorità di vigilanza qualora abbia motivo di ritenere che l'uso del sistema di IA ad alto rischio in conformità delle

[70] Nello specifico, per i sistemi di IA ad alto rischio, al capo III sezione 2 e, per i modelli di IA per finalità generali, al capo V sezioni 2 e 3.

[71] Vale a dire il regolamento che ha codificato le procedure per l'adozione delle norme vincolanti.

[72] Il considerando 17 del Reg. (UE) n. 1025/2012 espressamente prevede che «per rappresentanza di interessi sociali e parti sociali interessate alle attività di normazione europee si intendono in modo particolare le attività delle organizzazioni e delle parti che rappresentano i diritti fondamentali dei dipendenti e dei lavoratori, ad esempio i sindacati».

[73] Sul punto v. C. GALLESE, *La standardizzazione nell'AI Act*, in O. POLLICINO, F. DONATI, G. FINOCCHIARO, F. PAOLUCCI (a cura di), *La disciplina dell'Intelligenza Artificiale,* Giuffrè, Milano, 2025, p. 318 ss. Per completezza si rileva che, anche per far fronte all'assenza di adeguate norme armonizzate, la Commissione – nelle circoscritte ipotesi di cui all'art. 41 *dell'AI Act* – potrebbe adottare anche specifiche tecniche comuni. Si precisa, ad ogni modo, che il rispetto di siffatte prescrizioni farebbe sorgere – come chiarisce lo stesso art. 40 – una presunzione di liceità del rispetto delle previsioni, realizzando – di conseguenza – un bilanciamento meramente presuntivo degli interessi in gioco.

[74] Nei termini, ai sensi dell'art. 79, par. 1, di «prodotto che presenta un rischio» quale definito all'articolo 3, pt. 19, del regolamento (UE) 2019/1020, di cui si dirà a breve.

[75] Oltre che il fabbricante del prodotto, il rappresentante autorizzato, l'importatore o il distributore.

[76] Testualmente: «I deployer monitorano il funzionamento del sistema di IA ad alto rischio sulla base delle istruzioni per l'uso e, se del caso, informano i fornitori a tale riguardo conformemente all'articolo 72. Qualora abbiano motivo di ritenere che l'uso del sistema di IA ad alto rischio in conformità delle istruzioni possa comportare che il *sistema di IA*

istruzioni d'uso possa comportare un rischio nei termini di cui all'art. 79. Sempre l'art. 26, par. 5, sancisce difatti l'obbligo per il *deployer* di monitorare il funzionamento del sistema di IA ad alto rischio sulla base delle istruzioni per l'uso e sospenderne l'utilizzo qualora abbia motivo di ritenere che il prodotto possa presentare un rischio.

Quest'obbligo del *deployer* di monitorare il funzionamento del sistema di IA ad alto rischio (e di eventualmente sospenderne l'utilizzo) fa profilare una conformità "dinamica", che – come accennato – concerne gli sviluppi applicativi del sistema.

Per definire le caratteristiche del *sistema di IA che presenta un rischio* ai sensi dell'articolo 79, e quindi per definire i criteri di valutazione della conformità dinamica, è necessario leggere in combinato disposto l'art. 79 dell'*AI Act* e l'art. 3, pt. 19, del regolamento (UE) 2019/1020 (regolamento sulla vigilanza del mercato e sulla conformità dei prodotti, che lo stesso articolo 79, par. 1, richiama)[77]: il primo articolo indica i beni protetti, mentre il secondo individua il parametro di commisurazione per l'identificazione del sistema (nonché del rischio) in parola[78].

Ai sensi dell'art. 79, par. 1, infatti, «un sistema di IA che presenta un rischio» è inteso come un «prodotto che presenta un rischio» quale definito all'art. 3, pt. 19, del regolamento (UE) 2019/1020 «nella misura in cui presenta *rischi* per la salute o la sicurezza o *per i diritti fondamentali* delle persone»[79]. Affinché possa parlarsi di *sistema di IA che presenta un rischio* è necessario che per i beni protetti dalla disposizione dell'*AI Act,* vale a dire i diritti fondamentali della persona[80], si rilevi un potenziale pregiudizio, ai sensi dell'art. 3, pt. 19, «oltre quanto ritenuto ragionevole ed accettabile in relazione all'uso previsto del prodotto o nelle condizioni d'uso normali o ragionevolmente prevedibili»[81]. La norma richiamata integra, perciò, la disposizione dell'*AI Act* attraverso il riferimento alla soglia di ragionevolezza e accettabilità del rischio per i diritti fondamentali della persona, oltrepassata la quale il *deployer* è tenuto a sospendere l'utilizzo del

presenti un rischio ai sensi dell'articolo 79, paragrafo 1, i deployer ne informano, senza indebito ritardo, il fornitore o il distributore e la pertinente autorità di vigilanza del mercato e sospendono l'uso di tale sistema. Qualora abbiano individuato un incidente grave, i deployer ne informano immediatamente anche il fornitore, in primo luogo, e successivamente l'importatore o il distributore e le pertinenti autorità di vigilanza del mercato». Corsivo mio.

[77] Per comodità del lettore si riporta il testo dell'art. 3, pt. 19, del reg. (UE) 2019/1020, il quale definisce il «prodotto che presenta un rischio» come «un prodotto che potenzialmente potrebbe pregiudicare la salute e la sicurezza delle persone in generale, la salute e la sicurezza sul posto di lavoro, la protezione dei consumatori, l'ambiente e la sicurezza pubblica, nonché altri interessi pubblici tutelati dalla normativa di armonizzazione dell'Unione applicabile, oltre quanto ritenuto ragionevole ed accettabile in relazione all'uso previsto del prodotto o nelle condizioni d'uso normali o ragionevolmente prevedibili, incluse la durata di utilizzo e, se del caso, i requisiti relativi alla messa in servizio, all'installazione e alla manutenzione».

[78] Art. 79, par. 1, *AI Act*. Corsivo mio.

[79] Corsivo mio.

[80] Compresi i diritti fondamentali alla salute e alla sicurezza.

[81] Corsivo mio.

sistema (oltre che a informare il fornitore o il distributore[82] e la pertinente autorità di vigilanza del mercato).

Ne deriva che la conformità al regolamento, nella maggior parte dei casi – come si è detto – autocertificata dal produttore, non copre i concreti sviluppi legati all'utilizzo dell'IA[83]. In altri termini, la valutazione di conformità rappresenta l'elemento indispensabile per la sola commercializzazione del prodotto, giacché il *deployer*/datore di lavoro *è continuamente tenuto a valutare il concreto dispiegarsi del rischio* (sulla questione si tornerà, *infra* par. 6, anche per evidenziarne le ricadute sulla tecnica normativa dell'approccio basato sul rischio).

Tra gli ulteriori obblighi dei *deployer* previsti dall'art. 26 emergono, insieme al generale obbligo di fruire del servizio in conformità alle istruzioni d'uso (che, *ex* par. 1, richiede l'adozione di idonee misure tecniche e organizzative), quelli di: affidare la sorveglianza umana a persone competenti (par. 2); garantire che i dati di *input* siano pertinenti e sufficientemente rappresentativi in rapporto alla finalità del sistema (par. 4); segnalare incidenti gravi (tra i quali rientrano, *ex* art. 3, pt. 49, lett. *c*, la violazione degli obblighi a norma del diritto dell'Unione intesi a proteggere i diritti fondamentali); e, nel caso specifico dei *deployer* che sono datori di lavoro, informare i lavoratori interessati e i loro rappresentanti che saranno soggetti all'uso del sistema (par. 7).

4.4. (Segue) Il c.d. *FRIA*

Proseguendo nella disamina degli obblighi previsti dal regolamento, occorre soffermarsi sulla previsione secondo cui, ai sensi dell'art. 27, nell'ipotesi di servizi pubblici forniti tanto da un organismo di diritto pubblico quanto da un operatore privato (come i soggetti bancari o assicurativi), il *deployer* deve svolgere – prima della messa in uso del sistema – una valutazione d'impatto sui diritti fondamentali (c.d. *FRIA*, acronimo di *Fundamental Rights Impact Assessment*[84].

[82] Ossia, ai sensi dell'art. 4, pt. 7, «una persona fisica o giuridica nella catena di approvvigionamento, diversa dal fornitore o dall'importatore, che mette a disposizione un sistema di IA sul mercato dell'Unione».

[83] A medesime conclusioni, seppur attraverso differente argomentazione, giunge G. SMORTO, *Distribuzione del rischio,* cit., c. 217, il quale evidenzia che «In linea di principio, il regolamento si limita a disciplinare le condizioni di immissione sul mercato dei sistemi di IA. Mentre le conseguenze derivanti dagli impieghi di tali sistemi, una volta che siano immessi sul mercato, rimangono soggette alla disciplina di settore».

[84] Sul punto, cfr. A. MANTELERO, *The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template*, in *Computer Law & Security Review*, n. 54, 2024, p. 5 ss.; S. VILLANI, *Luci ed ombre degli strumenti di tutela dei diritti nell'architettura dell'AI Act basata sul rischio*, in F. LUNARDON, E. MENEGATTI (a cura di), *I nuovi confini del lavoro: la trasformazione digitale, Italian Labour Law e-Studies*, 2024, p. 145 ss., secondo la quale «sarebbe stato opportuno prevedere l'elaborazione di tali valutazioni … anche in relazione ai sistemi di IA a rischio limitato, visto il loro potenziale impatto sui diritti umani» (p. 146). Per un parallelismo con la valutazione di impatto sulla protezione dei dati personali *ex* art. 35 del GDPR, v. N. MINISCALCO, *L'*AI Act *nella prospettiva del diritto costituzionale: prime osservazioni*, in S. SCAGLIARINI, I. SENATORI (a cura di), *Lavoro, Impresa e Nuove Tecnologie, cit.*, pp. 24 e 25.

In dottrina si è però evidenziato come l'impiego del *FRIA* oltre il limitato ambito soggettivo tracciato dal regolamento sia comunque raccomandabile per prevenire conseguenze non solo sul piano risarcitorio ma anche su quello reputazionale e delle relazioni industriali[85]. Peraltro, come emerso dalle disposizioni esaminate nel precedente sottoparagrafo, il *deployer* è sempre gravato da un permanente obbligo di monitoraggio del sistema per valutarne l'impatto sui diritti fondamentali della persona, per quanto non si tratti di una valutazione d'impatto nei termini procedimentalizzati di cui all'art. 27, che richiederebbe – in particolare – la presa in considerazione dei « rischi specifici di danno che possono incidere sulle categorie di persone fisiche o sui gruppi di persone» interessati[86] e «le misure da adottare qualora tali rischi si concretizzino»[87].

In quest'ottica, il *FRIA*, lungi dall'essere un mero adempimento burocratico, può rappresentare un utile strumento organizzativo per facilitare il concreto adempimento degli obblighi del *deployer*.

4.5. Le pratiche vietate

Portando a termine la descrizione della menzionata quadripartizione dei sistemi di IA (*supra* par. 4), in un quarto gruppo, cui corrisponde un *rischio inaccettabile*, si individuano i sistemi vietati, ossia riconducibili alle pratiche di IA di cui all'art. 5 – unico articolo del capo II – le cui disposizioni sono già applicabili[88]. L'articolo bandisce – tra gli altri – i sistemi di riconoscimento delle emozioni di una persona fisica nell'ambito del luogo di lavoro (par. 1, lett. *f,* seppure con l'ampia eccezione d'uso dovuta a ragioni mediche o di sicurezza) e i sistemi di categorizzazione biometrica delle persone volti a trarre deduzioni o inferenze in merito a razza, opinioni politiche, appartenenza sindacale, convinzioni religiose o filosofiche, vita sessuale o orientamento sessuale (par. 1, lett. *g*). Questa classificazione rappresenta il risultato del lungo processo di gestazione del regolamento che, viceversa, nelle prime versioni presentava non pochi margini di ambiguità con riguardo alla possibilità di ricorrere, ad esempio, a sistemi di riconoscimento delle emozioni, i quali «rappresentano una delle modalità più insidiose (e subdole) di esercizio dei poteri datoriali»[89]. Proprio tale livello di rischio rappresenta, a ben vedere, la concretizzazione della necessaria

[85] A. MANTELERO, M. PERUZZI, *L'*AI Act *e la gestione del rischio, cit.* p. 534.
[86] Art. 27, par. 1, lett. *c.*
[87] Art. 27, par. 1, lett. *f.*
[88] Infatti, ai sensi dell'art. 113, l'applicazione dei Capi I e II decorre dal 2 febbraio 2025.
[89] Così L. TEBANO, *Intelligenza Artificiale e datore di lavoro*, *cit.*, p. 462, ciò perché «le emozioni sono fenomeni umani complessi che talvolta sfuggono al controllo dello stesso individuo che le prova e le manifesta (si pensi all'ansia o allo stress) … (e) perché la loro affidabilità è tutt'altro che pacifica». Sul punto, si segnala che la Commissione ha approvato, in data del 4 febbraio 2025, gli orientamenti sulle pratiche vietate (*Annex to the Communication to the Commission. Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act*), C(2025) 884 final), i quali, ad esempio, chiariscono che l'uso di webcam e di sistemi di riconoscimento vocale da parte di un call center per tracciare le emozioni dei dipendenti, come la rabbia, è vietato; mentre, se utilizzati solo per scopi di formazione personale, i sistemi di riconoscimento delle emozioni sono consentiti qualora i risultati non siano condivisi con i responsabili delle risorse umane e non possono influire, a titolo

interazione sistematica tra approccio basato sul rischio e principio di precauzione (cui si accennava *supra* par. 3). In altre parole, il legislatore europeo, vietando una serie di pratiche di IA, ha fatto ricorso al principio di precauzione in ragione dell'impossibilità di calcolare con precisione e in anticipo i potenziali gravi rischi legati all'utilizzo di questi sistemi. Ciò sembra confermare che l'approccio *risk-based* debba cedere il passo al principio di precauzione qualora si voglia limitare l'impiego di una tecnologia di cui non sia possibile valutare *ex ante* l'impatto.

Dunque, considerato che per il rischio inaccettabile la tecnica del *risk-based approach* si concretizza nel ricorso al principio di precauzione e che per i sistemi a rischio limitato la tutela è circoscritta alla trasparenza[90], nell'ambito dei sistemi ad alto rischio l'approccio assume un tratto peculiare. In questo caso, infatti, il rischio, oltre a rappresentare il fattore selettivo del regime giuridico[91], costituisce lo strumento di definizione "sartoriale" dell'intelaiatura delle tutele a protezione dei diritti fondamentali relative allo specifico sistema di IA, tanto che si potrebbe parlare di "approccio basato sul rischio in senso stretto".

5. Tipologia del rischio

È opportuno, a questo punto, porre l'accento sul concetto stesso di rischio, dato che nelle pagine precedenti il termine è stato utilizzato con una pluralità di significati e si è accompagnato ad una serie di aggettivazioni, le quali vanno oltre la riportata quadripartizione classificatoria dei sistemi di IA[92].

In primo luogo, è evidente che quando si parla di approccio basato sul rischio si fa riferimento alla tecnica di regolazione prescindendo dalla connotazione giuridica che al *rischio* può essere attribuita da questo o quell'atto. Nell'*AI Act*, come si è visto, il legislatore ha definito sia il *rischio* (in generale) sia la specifica ipotesi di *rischio sistemico*.

Il *rischio* (in generale) *ex* art. 3, pt. 2, viene definito come un fattore probabilistico frutto della combinazione di due elementi: la probabilità del verificarsi di un danno e la gravità del danno stesso.

Una particolare declinazione di questo concetto si ritrova nella disciplina dei sistemi di IA ad alto rischio, per la cui qualificazione l'art. 6, par. 3, richiede la sussistenza di un *rischio significativo di danno* (*supra* par. 4.1) per la salute, la sicurezza o i diritti fondamentali delle persone fisiche.

Diverso è il caso del *rischio sistemico* (*ex* art. 3, pt. 65), che richiede una trattazione autonoma sia per la peculiare natura sia per la specifica disciplina, avendo il regolamento previsto regole *ad hoc* per i prodotti

esemplificativo, sulla valutazione e sulla promozione della persona formata, a condizione che il divieto non sia aggirato e che l'uso del sistema di riconoscimento delle emozioni non abbia alcun impatto sul rapporto di lavoro (p. 85).

[90] Considerato anche che per il rischio minimo non sono necessarie specifiche forme di tutela.

[91] Sulla scia della definizione regolamentare del cons. 26, *supra* par. 1.

[92] Sulla non neutralità della nozione di rischio v. M. FAIOLI, *Assessing Risks and Liabilities of AI-Powered Robots in the Workplace. An EU-US Comparison*, in *Diritto della Sicurezza sul Lavoro*, n. 1, 2025, p. 87.

che costituiscono la fonte del rischio in parola[93]. Come anticipato[94], il *rischio sistemico* consiste in un «rischio specifico» che è potenziale conseguenza della capacità di impatto elevato dei *modelli di IA per finalità generali*[95], i quali per via della loro, appunto, «generalità» e della capacità di «svolgere con competenza un'ampia gamma di compiti distinti»[96] presentano sfide regolatorie specifiche. Tra i modelli di IA per finalità generali rientrano – per espressa previsione dell'*AI Act* [97] – i modelli di IA generativa, come il modello *Generative Pre-trained Transformer* (*GPT*), sul quale si sono basate, ad esempio, le diverse versioni di *Chat GPT* sviluppate da *OpenAI*. Il rischio sistemico si sostanzia in un «impatto significativo sul mercato dell'Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso»[98]. Per quanto, come si è detto, il rischio sistemico presenti delle spiccate specificità, il riferimento agli *effetti negativi effettivi* potrebbe corroborare l'idea che il complessivo programma normativo tolleri le conseguenze pregiudizievoli derivanti dall'utilizzo dei "prodotti intelligenti" messi in commercio in conformità all'ordinamento dell'Unione europea, problematica sulla quale ci si soffermerà nel prossimo paragrafo.

6. Rischio, danno e *autoreferenzialità nel bilanciamento* degli interessi

Dall'analisi delle molteplici definizioni di *rischio* ("senza aggettivi", *significativo, sistemico*) presenti nel regolamento sull'IA è venuta in rilievo una questione cruciale: la relazione tra rischio e danno.

La scelta dell'*AI Act* di definire il rischio in termini di probabilità del verificarsi di un danno porta con sé il pericolo di non tenere ben distinti questi due elementi, con la conseguenza che l'intero approccio regolamentare potrebbe apparire "*harm-based*", ossia graduato in relazione alle ricadute pregiudizievoli sui diritti fondamentali. In altre parole, differentemente dall'impostazione adottata nel GDPR – secondo una concezione *right-based* – che poggia sui diritti e sulla loro rilevanza, improntare le tutele al rischio di danno potrebbe far pensare che, una volta rispettati i requisiti di conformità previsti dal regolamento, l'effettivo

[93] Il regolamento dedica specifiche regole ai modelli di IA per finalità generali (capo V, art. 51 ss.) e pone particolare attenzione ai modelli con rischio sistemico, prevedendo specifici e ulteriori obblighi in capo ai loro fornitori (capo V, sez. 3, art. 55), tra i quali quello di valutare e attenuare i possibili rischi sistemici (art. 55, par. 1, lett. *b*). La conformità a questi obblighi, in attesa di una norma armonizzata alla quale si rinvia, può essere dimostrata attraverso i codici di buone pratiche (art. 55, par. 2). In generale, il potere di monitorare e supervisionare la conformità è affidato all'Ufficio per l'IA (Art. 75, par. 1).

[94] La definizione è stata già riportata nella nota 17.

[95] Sul rapporto tra sistema e modello il cons. 97 chiarisce che: «sebbene i modelli di IA siano componenti essenziali dei sistemi di IA, essi non costituiscono di per sé sistemi di IA». Corsivo aggiunto.

[96] V. art. 3, pt. 63.

[97] Ai sensi del considerando 99: «I grandi modelli di IA generativi sono un tipico esempio di modello di IA per finalità generali, dato che consentono una generazione flessibile di contenuti, ad esempio sotto forma di testo, audio, immagini o video, che possono prontamente rispondere a un'ampia gamma di compiti distinti». Anche al considerando 105 si legge: «I modelli di IA per finalità generali, in particolare i grandi modelli di IA generativa, in grado di generare testo, immagini e altri contenuti, presentano opportunità di innovazione uniche, ma anche sfide per artisti, autori e altri creatori e per le modalità con cui i loro contenuti creativi sono creati, distribuiti, utilizzati e fruiti».

[98] Così art. 3, pt. 65.

concretizzarsi del danno sia da "accettare" in quanto conseguenza di un'operazione di bilanciamento connaturata nell'approccio *risk-based* delineato dal legislatore europeo. Si potrebbe, perciò, ipotizzare che un datore di lavoro che utilizzi un sistema di IA ad alto rischio – necessariamente già valutato conforme al regolamento – non possa essere chiamato a rispondere di conseguenze pregiudizievoli arrecate ai lavoratori dalla strumentazione "intelligente" qualora l'impiego sia stato effettuato nel rispetto delle istruzioni d'uso, in quanto il pregiudizio sarebbe il residuo del bilanciamento effettuato dal regolamento con le esigenze del mercato unico, il cui miglioramento rappresenta uno scopo dichiarato sin dall'*incipit* dell'art. 1.

Questa lettura, come si proverà ad evidenziare, non si sposa però con un'interpretazione sistematica che tiene conto del complessivo impianto normativo dell'*AI Act* e della sua collocazione nell'ordinamento dell'Unione.

Come questione preliminare rispetto a ogni considerazione sul rapporto tra il rischio di danno e il danno medesimo si pone il tema del grado di rischio tollerato dal regolamento. In dottrina è stato evidenziato come l'*AI Act* non miri all'eliminazione dei rischi[99]; d'altronde, pur volendo ammettere la possibilità di eliminare il rischio, nel regolamento l'unico riferimento all'eliminazione si registra solo come alternativa alla riduzione all'art. 9, par. 5, ove «Nell'individuare le misure di gestione dei rischi più appropriate» si considera necessario garantire «l'eliminazione o la riduzione dei rischi individuati».

Appurato che il regolamento tollera il rischio residuo per l'immissione del prodotto sul mercato, bisogna verificare se questa tolleranza si estenda (ed eventualmente in che misura) anche alle fasi successive alla commercializzazione, potendo in tal modo mettere in pericolo il sistema di tutele giuslavoristiche dal carattere prescrittivo[100]. L'elemento chiave per rispondere a quest'interrogativo può rinvenirsi nell'obbligo di monitoraggio gravante sul *deployer*. Come si è detto (*supra* par. 4.3), attraverso il combinato degli artt. 26 e 79, l'*AI Act* obbliga il *deployer* a monitorare il funzionamento del sistema di IA ad alto rischio e sospenderne l'utilizzo qualora abbia motivo di ritenere che il prodotto possa pregiudicare la salute, la sicurezza o i diritti fondamentali delle persone *oltre quanto ritenuto ragionevole e accettabile*. L'obbligo in parola, quindi, si pone come argine *ultimo* allo straripamento del *rischio residuo*, il cui governo viene affidato al *deployer*.

Si tratta in effetti di un'operazione di bilanciamento che però non è "strutturata" e, perciò, presenta confini incerti. Ciò che lascia perplessi è l'assenza di elementi che possano indirizzare il *deployer*/datore di lavoro, diversamente da quanto riscontrabile – ad esempio – nella disciplina degli accomodamenti

[99] C. LAZZARI, P. PASCUCCI, *Sistemi di IA, salute e sicurezza, cit.*, p. 595; G. GAUDIO, V*alutazioni d'impatto e management algoritmico, cit.*, p. 543 ss., il quale evidenzia come – invece – l'eliminazione del rischio sia un obiettivo della direttiva piattaforme.

[100] Preoccupazione sollevata da P. LOI, *Il rischio proporzionato*, *cit.*, 295 ss.; A. ALOISI, V. DE STEFANO, *Between risk mitigation and labour rights enforcement, cit.*, p. 283 ss.

ragionevoli[101]. Non si tratta – naturalmente – di richiedere indicazioni precise e definite, ma binari e procedure che conferiscano maggiore solidità all'operazione, almeno prima di arrivare a un'ipotesi di contenzioso.

Tirando le fila del ragionamento sviluppato, se l'ordinamento non tollera che l'utilizzo di un sistema di intelligenza artificiale – qualificato dal fornitore come ad alto rischio – possa produrre un *rischio "non ragionevole"* (nei termini di cui sopra), non può di certo tollerare le conseguenze pregiudizievoli che di questo *rischio* rappresentano la concreta e non più potenziale manifestazione. Pertanto, se vi è un soggetto – quale il datore di lavoro – deputato a vigilare sul non travalicamento della soglia di rischio accettabile per la salute, per la sicurezza e (in generale) per i diritti fondamentali della persona, la sua condotta omissiva avrà rilievo sul piano della responsabilità civile[102]. Ponendo l'accento sugli obblighi del *deployer* la lettura della tecnica dell'approccio basato sul rischio si rivela dunque "*right-based*" (precisamente "*fundamental right-based*") impedendo che il diritto venga svuotato. In effetti è questa l'essenza dell'operazione di bilanciamento dei diritti che fin qui ha trovato, come si sa, compiuto riscontro normativo nell'art. 52 della Carta dei diritti fondamentali dell'Unione europea[103], il quale consente limitazioni dei diritti fondamentali nel rispetto del principio di proporzionalità[104]. Limitazioni legittime purché siano previste dalla legge e siano necessarie e rispondenti effettivamente a finalità di interesse

[101] L'art. 5 *bis,* l. 5 febbraio 1992, n. 104, introdotto con dall'art. 17 del d.lgs. 3 maggio 2024, n. 62, consente infatti alla persona con disabilità di partecipare al procedimento relativo all'individuazione dell'accomodamento ragionevole (co. 4), potendone anche «richiedere, con apposita istanza scritta, … l'adozione» (co. 3), proposta la cui possibilità di accoglimento deve essere previamente verificata nella valutazione dell'istanza di accomodamento (co. 6). Come è stato evidenziato, si tratta di una procedimentalizzazione del potere del datore di lavoro con il previo coinvolgimento del beneficiario: cfr. D. TARDIVO, *L'inclusione lavorativa della persona con disabilità: tecniche e limiti*, Giappichelli, Torino, 2024, p. 156.

[102] Sembrerebbe così delinearsi anche un ampliamento della sfera dei soggetti responsabili, tema sul quale invita a riflettere L. ZOPPOLI, *Il diritto del lavoro dopo l'avvento dell'intelligenza artificiale, cit*., p. 9, che si chiede: «Se esistono corresponsabilità soggettive *by design* o *by monitoring* o *by control*, ampliare la sfera dei soggetti responsabili non può essere preferibile anche in chiave di deterrenza?».

[103] V. M. DELFINO, *Article 52 CFREU*, in E. ALES, M. BELL, O. DEINERT, S. ROBIN-OLIVIER (eds.), *International and European Labour Law*, Nomos, Baden-Baden, 2018, p. 239 ss. In relazione all'ambito di applicazione della Carta si ricorda che, ai sensi dell'art. 51 CDFUE, le disposizioni della Carta «si applicano alle istituzioni, organi e organismi dell'Unione …, come pure agli Stati membri esclusivamente nell'attuazione del diritto dell'Unione»; sul punto cfr. M. DELFINO, *Article 51 CFREU*, in E. ALES, M. BELL, O. DEINERT, S. ROBIN-OLIVIER (eds.), *International and European Labour Law, cit*., (II edizione in corso di pubblicazione, par. IV), secondo il quale l'articolo non definisce l'ambito soggettivo, bensì quello materiale e, nello specifico «the scope of application of the Charter, in compliance with the Union's competences, which the Charter cannot modify or expand».

[104] Sul bilanciamento tra i diritti fondamentali in relazione all'*AI Act* cfr. P. LOI, *Il rischio proporzionato, cit*., pp. 253-255. In generale sui principi di ragionevolezza e proporzionalità, *ex multis*, G. SCACCIA, *Proporzionalità e bilanciamento tra diritti nella giurisprudenza delle corti europee*, in *Rivista AIC*, n. 3, 2017, p. 7 ss.; in ambito giuslavoristico, v. P. LOI. *Il principio di ragionevolezza e proporzionalità nel diritto del lavoro*, Giappichelli, Torino, 2016. In relazione all'importanza del compiuto bilanciamento degli interessi sin dalla stesura dell'atto legislativo cfr. A. ZOPPOLI, *Prospettiva rimediale, fattispecie, sistema*, in *Le tecniche di tutela nel diritto del lavoro. Atti giornate di studio AIDLaSS*, Torino, 16-17 giugno 2022, La Tribuna, Milano, 2023, p. 147: «La legge (…) definisce la "fisionomia" dell'assetto degli interessi in gioco, quindi il loro primo bilanciamento e la loro prima "mediazione", secondo il circuito fondativo dell'ordinamento democratico».

generale riconosciute dall'Unione o all'esigenza di proteggere i diritti e le libertà altrui e sempre che dei diritti sia rispettato il contenuto essenziale (art. 52, par. 1, CDFUE).

Ricapitolando, con l'*AI Act* il decisore politico si è focalizzato sull'individuazione delle pratiche vietate e delle aree critiche e degli specifici settori ad alto rischio. La regolamentazione dei sistemi di IA ad alto rischio si caratterizza per l'elencazione di obiettivi e indicazioni generali che lasciano un significativo margine di discrezionalità tecnica nella valutazione di conformità (necessaria per l'immissione del prodotto sul mercato) che è affidata (nella maggior parte dei casi) all'autocertificazione del produttore. Siffatta attestazione non può strutturalmente dar luogo a un adeguato contemperamento tra i diritti in gioco per via dell'ampia discrezionalità e della corrispondenza tra il certificatore e il soggetto che dalla commercializzazione trae profitto. Ne consegue che la certificazione di conformità comporta la sola possibilità di immettere il prodotto sul mercato, come evidenziato (*supra* par. 4.2) in relazione all'*autoreferenzialità certificatoria* del fornitore.

Per quanto concerne gli sviluppi applicativi, vale a dire la *conformità dinamica*, lo stesso *AI Act* obbliga il *deployer* a monitorare il sistema e a sospenderne l'utilizzo qualora abbia motivo di ritenere che il prodotto possa pregiudicare i diritti fondamentali della persona *oltre quanto ritenuto ragionevole e accettabile* (v. *supra* par. 4.3). Neanche in questa circostanza il criterio del contemperamento viene chiaramente definito né dal regolamento né dalla valutazione di conformità e il datore di lavoro si ritrova a dover vigilare, con le lenti della *ragionevolezza*, su un'imprecisata frontiera che separa ciò che è danno da ciò che è legittima compressione dei diritti fondamentali della persona. Si ripropone, così, in una nuova veste, il tema dell'autoreferenzialità, questa volta del *deployer*/datore di lavoro in relazione alla scelta di sospendere l'utilizzo del sistema di IA. In altre parole, il limite delineato è incentrato – senza che siano forniti elementi "tecnici" – sul bilanciamento degli interessi, la cui ponderazione, in assenza di una soddisfacente valutazione legislativa, viene rimessa alla valutazione datoriale. La tecnica adottata è quindi di per sé foriera di un margine di scelta del soggetto deputato al bilanciamento, ossia all'individuazione del limite ragionevole e accettabile per il rischio che – in prima battuta – è rimesso al datore. Emerge, quindi, la concreta esigenza di ridurre l'incertezza, non positiva per le relazioni di lavoro, che caratterizza l'attuale quadro normativo. Una soluzione potrebbe essere rappresentata da percorsi di proceduralizzazione che possano fornire criteri operativi condivisi. In assenza di simili strumenti, nell'indeterminatezza dell'impostazione normativa, è verosimile che si ricorra al giudice, perpetuando il clima di incertezza interpretativa. La descritta *autoreferenzialità nel bilanciamento degli interessi* risulta dunque intrinsecamente problematica e apre ad altre questioni relative agli sviluppi una simile tecnica nonché all'inquadramento sistematico del sotteso obbligo datoriale di monitorare ed eventualmente sospendere l'utilizzo del sistema di IA ad alto rischio, profili che si propongono come terreno di ricerca per la dottrina giuslavoristica.