Aspen Institute Italia

OSSERVATORIO PERMANENTE SULL'ADOZIONE E L'INTEGRAZIONE DELLA INTELLIGENZA ARTIFICIALE (IA^2)

RAPPORTO INTELLIGENZA ARTIFICIALE 2025

In collaborazione con

INTESA SANPAOLO

Aspen Institute Italia

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

INDICE

Il Rapporto è stato chiuso con le informazioni dati disponibili al 3 giugno 2025.

Cura e revisione del testo Simonetta Savona, Aspen Institute Italia.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

EXECUTIVE SUMMARY

L'intelligenza artificiale è sempre più al centro di una competizione globale. **Stati Uniti, Cina ed Europa stanno adottando approcci diversi sull'IA**, bilanciando investimenti pubblici e privati in funzione delle rispettive priorità. Gli Stati Uniti accelerano tramite consorzi privati come Stargate; la Cina punta sulla centralizzazione statale per ottenere autosufficienza e primato tecnologico entro il 2030; l'Europa ha puntato sul fronte regolatorio e ora sta investendo sul sostegno al proprio tessuto industriale. **L'IA non è più solo una tecnologia, ma una leva strategica.**

La struttura a cinque layer dell'ecosistema IA – *hardware*, *cloud computing*, dati, *foundation models*, applicazioni – definisce una nuova mappa globale. Gli Stati Uniti si focalizzano principalmente sui **primi due livelli (*hardware* e *cloud*)**, mentre la Cina avanza su modelli prevalentemente chiusi e sovranità sui

dati. Quanto all'UE, sembra prospettarsi una "terza via", fondata sulla normazione, sull'etica, sulla sicurezza e sulla tutela dei diritti fondamentali alla base della tradizione democratica del Vecchio Continente, ma anche su innovazione e sostenibilità attraverso una spinta abilitativa, anche in termini tecnologici e infrastrutturali, di sprone agli investimenti e di rafforzamento delle competenze, per realizzare un'IA affidabile per un'Europa più competitiva, che possa ricoprire una posizione di primo piano nel settore e possa esprimere al meglio il proprio potenziale nel settore con utilizzi originali, creativi e nuovi.

Nonostante l'entusiasmo, l'impatto economico dell'IA solleva interrogativi. Secondo Acemoğlu, **i benefici sull'aumento di produttività potrebbero essere limitati nel breve termine**, con effetti stimati inferiori all'1,6% sul PIL in dieci anni. Tuttavia, l'IA rimane una tecnologia "*general purpose*" capace di trasformare il rapporto tra capitale e lavoro, grazie all'automazione cognitiva. **L'IA può sbloccare nuovi margini di produttività, ma solo se integrata in modo intelligente nei sistemi economici.**

Il mercato del lavoro è il primo a essere trasformato. Non si osserva una distruzione netta dei posti di lavoro, ma **una riconfigurazione profonda delle competenze richieste**, con effetti asimmetrici: **emergono nuove professioni ibride**, mentre le mansioni ripetitive e poco qualificate possono risultare più esposte, laddove i costi della tecnologia siano inferiori e non necessitino di destrezza fine. È vero però anche il contrario, le professioni più esposte sono quelle di natura cognitiva, come specificato all'interno del rapporto, sia pure non in questa estrema sintesi. Il rischio principale è il *mismatch* temporale tra skill disponibili e skill richieste.

L'adozione dell'IA senza politiche formative strutturate però può amplificare le disuguaglianze sociali.

Per affrontare queste transizioni, il *lifelong learning* è imprescindibile. Paesi come Singapore (con *SkillsFuture*) o Finlandia stanno **sperimentando ecosistemi formativi distribuiti**, capaci di abilitare la formazione continua. Anche in Italia sono attivi strumenti come il Fondo Nuove Competenze e gli enti bilaterali, serve però **una strategia integrata che possa consentire alle organizzazioni di diventare anche delle *learning companies***. Le scuole, inoltre, devono passare da luoghi di trasmissione passiva ad **ambienti di stimolazione del pensiero critico e dell'intelligenza umana "aumentata"**.

Nei settori applicativi, l'IA sta già trasformando interi comparti. In sanità riduce tempi e costi nella scoperta di nuove molecole e abilita una medicina personalizzata. Nella finanza **potenzia le strategie antiriciclaggio e abilita strumenti finanziari su misura**. Nei beni culturali consente la digitalizzazione del patrimonio, ma **solleva interrogativi sull'autenticità delle riproduzioni e sul rischio di una fruizione superficiale**. Nel settore sportivo, grazie a *wearable* e *analytics* predittivi, **ottimizza le prestazioni e previene gli infortuni**.

L'IA sta diventando un alleato cruciale per la sostenibilità ambientale e la transizione ecologica. I modelli predittivi migliorano la gestione delle risorse naturali, il monitoraggio ambientale e l'efficienza energetica. Dall'agricoltura di precisione alla manutenzione predittiva di infrastrutture critiche, **l'IA può consentire una riduzione significativa delle emissioni e degli sprechi.** Tuttavia, permangono contraddizioni: i consumi energetici dei grandi modelli e dei data center pongono sfide inedite alla sostenibilità. **L'adozione dell'IA richiede una**

valutazione del ciclo di vita energetico e l'integrazione con fonti rinnovabili. Una politica industriale sostenibile deve coniugare innovazione digitale e transizione verde.

A livello etico e sociale, **l'IA sfida i principi fondamentali di equità, trasparenza e responsabilità.** Le sfide riguardano il rischio di discriminazione algoritmica, l'erosione della *privacy*, l'opacità decisionale e il possibile impoverimento cognitivo. Il concetto di "Sistema 0" richiama la necessità di **non delegare il pensiero critico ai sistemi "intelligenti"**, mantenendo l'uomo al centro dell'interpretazione e della decisione. **Serve una nuova cultura dell'interazione uomo-macchina, basata su consapevolezza, formazione e vigilanza.**

Le dimensioni etico-sociali dell'IA richiedono una *governance* multilivello, capace di integrare competenze tecniche, riflessioni filosofiche, norme giuridiche e rappresentanza democratica. Senza un nuovo patto sociale e culturale, la tecnologia rischia di acuire le fratture esistenti e di disancorarsi dai valori fondativi delle democrazie liberali. È urgente includere nella progettazione dell'IA aspetti legati all'equità, all'inclusione e alla diversità, promuovendo **un'intelligenza artificiale umanistica e pluralista.**

Sul piano normativo, **il mondo si muove con approcci diversi.** Gli Stati Uniti si affidano a principi di autoregolazione, la Cina a un controllo diretto e centralizzato, mentre l'Unione Europea ha adottato l'AI Act come tentativo di **definire uno standard globale basato su trasparenza e tutela dei diritti e su responsabilità e innovazione.**

Anche al fine di evitare l'incertezza connessa alla frammentazione normativa del settore, una cooperazione

internazionale - almeno tra le democrazie liberali - è sempre più necessaria per **definire cornici comuni su standard, *auditing*, responsabilità e sicurezza.** L'assenza di una convergenza tra normative rischia di generare barriere, rallentando lo sviluppo di applicazioni *cross-border* e accentuando gli squilibri globali.

Sul fronte tecnico, il 2024 ha segnato il consolidamento di tre direttrici: **IA generativa, IA collaborativa e tecnologie per una IA *privacy-preserving***. I *Large Language Models* si diffondono nell'automazione dei contenuti, l'IA collaborativa abilita interazioni uomo-macchina in sanità e manifattura, mentre tecnologie come il *federated learning* consentono di **usare dati sensibili senza comprometterne la *privacy***.

Se si guarda al futuro, emerge il tema dell'Intelligenza Artificiale Generale (AGI). **Sistemi con capacità cognitive avanzate potrebbero generare comportamenti imprevisti**, sfidando i tradizionali strumenti di controllo e allineamento etico. L'assenza di una *governance* multilivello condivisa potrebbe amplificare i rischi. **Senza cooperazione strutturata a livello globale, l'AGI rischia di evolvere in modo opaco e non democratico.**

Questo rapporto vuole offrire uno strumento di comprensione critica e sistemica dell'IA con lo scopo precipuo di comprendere come essa è stata adottata nella realtà italiana dell'IA. **Governare l'intelligenza artificiale significa coniugare tecnologia, etica, competenze e regolazione**, costruendo una traiettoria che sia al tempo stesso competitiva, sostenibile, inclusiva e orientata al miglioramento del benessere delle persone e della società nel suo complesso. Per farlo, serve una visione politica ambiziosa, capace di guidare l'innovazione nel rispetto della dignità umana.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

INTRODUZIONE

L'intelligenza artificiale rappresenta oggi uno dei principali motori del cambiamento economico, sociale e tecnologico a livello globale. A testimonianza di ciò, il Presidente degli Stati Uniti, Donald Trump, ha annunciato nel gennaio 2025 la creazione di "*Stargate*", una *joint venture* tra OpenAI, Oracle e SoftBank[1], con l'obiettivo di investire fino a cinquecento miliardi di dollari nello sviluppo di infrastrutture per l'IA nei prossimi quattro anni.

[1] Cfr. https://edition.cnn.com/2025/01/21/tech/openai-oracle-softbank-trump-ai-investment/index.html.

Parallelamente, l'Unione Europea ha lanciato l'iniziativa InvestAI[2], destinata a mobilitare duecento miliardi di euro per promuovere la ricerca e l'innovazione nell'intelligenza artificiale, con l'obiettivo di accelerare lo sviluppo della tecnologia nel Vecchio Continente. Questi ingenti investimenti evidenziano come la corsa all'IA sia diventata una questione non solo tecnica, ma anche strategica a livello industriale, sociale e geopolitico.

Ovviamente, in parallelo agli entusiasmi più diffusi, emergono anche voci che invitano a un ridimensionamento delle aspettative legate all'IA. Tra queste si distingue quella dell'economista Daron Acemoğlu[3] del MIT, il quale sottolinea come - nonostante gli investimenti massicci - l'impatto economico reale dell'IA potrebbe essere meno incisivo del previsto. Secondo Acemoğlu, infatti, nel prossimo decennio l'intelligenza artificiale contribuirà a un aumento modesto del PIL statunitense, compreso tra l'1,1% e l'1,6%, con un incremento annuale della produttività di circa lo 0,05%, suggerendo così una valutazione più cauta rispetto alle previsioni più ottimistiche.

[2] Cfr. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467.

[3] Cfr. https://economics.mit.edu/news/daron-acemoglu-what-do-we-know-about-economics-ai.

Cosa riservi il futuro è difficile a sapersi, però, se è arduo credere che l'intelligenza artificiale possa rappresentare la soluzione definitiva a tutti i problemi globali, è altrettanto difficile pensare che l'interesse internazionale per questa tecnologia sia un mero abbaglio collettivo, allora negare che le modalità di interazione con la tecnologia stiano evolvendo sarebbe un errore.

In tale complesso contesto nasce questo Rapporto. Nasce dalla consapevolezza che adozione e integrazione dell'IA non costituiscono più una semplice opzione tecnologica, bensì una necessità strategica imprescindibile per governi, imprese e istituzioni.

La diffusione dell'intelligenza artificiale, infatti, non riguarda solo un incremento dell'efficienza operativa o una riduzione dei costi; piuttosto, essa implica una trasformazione profonda e strutturale dell'economia e della società, che coinvolge temi quali la sostenibilità, l'etica, la sicurezza e le politiche normative.

Il presente *Rapporto dell'Osservatorio Permanente sull'Adozione e l'Integrazione dell'IA 2025* segue il primo, pubblicato nel 2024 da

Aspen Institute Italia[4]. Mentre il precedente rapporto (2024) si concentrava principalmente sull'analisi delle potenzialità emergenti dell'intelligenza artificiale e sulle sue applicazioni iniziali, qui si approfondirà l'evoluzione dell'adozione dell'IA nell'ultimo anno, evidenziando le tendenze attuali, le sfide emergenti e le strategie adottate a livello globale.

Il Rapporto è strutturato in diverse sezioni per offrire una visione completa dell'argomento. Dopo l'introduzione, la prima sezione fornisce un'analisi delle tendenze globali nell'adozione dell'IA, seguita da una valutazione delle implicazioni economiche e sociali. A seguire, sono esaminati casi studio specifici in settori chiave come la sanità, la finanza, il patrimonio culturale, l'agricoltura e lo sport. Infine, il rapporto affronta le questioni etiche e normative, concludendo con raccomandazioni strategiche per i decisori politici e gli stakeholder coinvolti.

Attraverso un'analisi multidimensionale e interdisciplinare, verranno esplorate le dinamiche settoriali in aree strategiche quali la sanità, la finanza, il *cultural heritage*, l'agricoltura, lo sport,

[4] ASPEN INSTITUTE ITALIA (a cura di), *Rapporto Intelligenza artificiale 2024*, Roma 2024, cfr. https://www.aspeninstitute.it/rapporto-intelligenza-artificiale-2024/.

le telecomunicazioni, il manifatturiero e la difesa e sicurezza, evidenziando sia le straordinarie opportunità offerte dall'IA sia i rischi connessi a una sua integrazione massiva.

Sarà inoltre posta particolare attenzione alle problematiche etico-sociali e alle politiche regolatorie, fattori essenziali per garantire uno sviluppo equilibrato e sostenibile dell'intelligenza artificiale, capace di generare benefici diffusi senza sacrificare sicurezza e inclusione sociale.

L'obiettivo finale del Rapporto è fornire una piattaforma di conoscenza e confronto che favorisca una comprensione strategica dell'IA, affiancando decisori, stakeholder e cittadini nella costruzione consapevole e responsabile di un futuro sempre più modellato dalle tecnologie intelligenti.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

I.

ADOZIONE DELL'IA: TENDENZE GLOBALI E REGIONALI

Con il contributo di

Alessandro Piccioni, Paolo Bellini, Edoardo degli Innocenti

L'adozione dell'intelligenza artificiale è oggi uno dei fenomeni più rilevanti nella trasformazione dei sistemi economici, politici e sociali a livello globale. Il capitolo esplora le principali tendenze di diffusione dell'IA a livello mondiale e regionale, analizzando le forze che ne guidano la messa in opera, le asimmetrie nei modelli di sviluppo e le implicazioni strategiche nei diversi settori industriali. A partire da una prospettiva

macroeconomica, si approfondisce qui il ruolo dell'IA nella ridefinizione dei paradigmi di competitività e nella trasformazione delle catene del valore.

Sono inoltre qui esaminate le diverse traiettorie di adozione nelle grandi aree geopolitiche - Stati Uniti, Cina, Unione Europea - con particolare attenzione alla complementarità tra dinamiche di mercato e intervento pubblico. Infine, si analizzano i principali rischi e opportunità legati a una diffusione massiva dell'IA, con focus specifici sull'impatto nel mercato del lavoro, sulla produttività e sulla sostenibilità dei modelli di crescita.

I.1 Impatto dell'IA sulle dinamiche economiche globali

L'adozione dell'intelligenza artificiale nell'economia globale non può essere ridotta a una mera questione di efficienza operativa e riduzione dei costi. Al contrario, essa rappresenta un fenomeno complesso, radicato in trasformazioni strutturali dell'economia, dinamiche tecnologiche di portata sistemica e tendenze settoriali che, negli ultimi anni, hanno ridefinito i paradigmi della produttività, della competitività e della *governance* economica. Di fatto la sua diffusione non è il risultato di un semplice progresso tecnologico, ma la risposta a pressioni

economiche e geopolitiche che spingono governi e imprese a ripensare le modalità di allocazione del capitale[5], l'organizzazione del lavoro e la creazione di vantaggi competitivi sostenibili.

Qui si analizzano le forze trainanti che hanno determinato l'adozione dell'intelligenza artificiale nell'economia globale, con un focus su otto settori strategici: la sanità, il comparto finanziario e bancario, il settore manifatturiero e dell'automazione industriale, il *cultural heritage*, l'agricoltura, lo sport, le telecomunicazioni e la difesa e sicurezza.

L'obiettivo è offrire una prospettiva macroeconomica che, oltre a contestualizzare le tendenze osservate nel recente passato, delinei le dinamiche emergenti che influenzeranno la traiettoria evolutiva dell'intelligenza artificiale nei prossimi anni.

[5] COMMISSIONE EUROPEA, "L'UE lancia l'iniziativa InvestAI per mobilitare 200 miliardi di € di investimenti nell'intelligenza artificiale", comunicato stampa, 11 febbraio 2025, https://digital-strategy.ec.europa.eu/it/news/eu-launches-investai-initiative-mobilise-eu200-billion-investment-artificial-intelligence.

I.1.2 FORZE TRAINANTI DELL'ADOZIONE DELL'IA NELL'ECONOMIA GLOBALE

L'intelligenza artificiale non si limita all'elaborazione di grandi volumi di dati, ma ne consente la trasformazione in un asset economico strategico. A differenza delle risorse tradizionali, come il petrolio, i dati presentano rendimenti marginali crescenti: il loro valore non si esaurisce con l'uso, bensì aumenta progressivamente attraverso l'addestramento dei modelli e l'apprendimento iterativo.

È una caratteristica tale da rendere l'accesso e l'utilizzo dei dati un elemento centrale nella competizione economica e geopolitica globale. Governi e imprese si confrontano per il controllo e la valorizzazione dei dati, influenzando le politiche commerciali, le normative sulla concorrenza e le strategie di sicurezza nazionale. **L'asimmetria nell'accesso ai dati determina vantaggi competitivi significativi**, favorendo l'emergere di nuovi paradigmi economici basati sulla capacità di integrare, analizzare e monetizzare informazioni in tempo reale. Di conseguenza, la gestione dei dati non è più un semplice fattore tecnico, ma una

leva fondamentale nella ridefinizione delle dinamiche di potere economico e tecnologico a livello globale.

Oltre al tema della competizione per il controllo dei dati, è utile analizzare il mercato dell'intelligenza artificiale nella sua articolazione tecnica. Un utile modello concettuale è quello dei cinque strati (*layers*) dell'ecosistema IA, che consente di comprendere l'intera catena del valore tecnologico e le interdipendenze strategiche tra attori pubblici e privati.

I cinque layer sono: *Hardware, Cloud Computing, Training Data, Foundation Models* e *AI Applications.*

Il *layer* ***hardware*** comprende semiconduttori e chip specializzati per l'intelligenza artificiale, come le GPU prodotte da NVIDIA o AMD, o le TPU sviluppate da Google. Tali componenti sono essenziali per le fasi di addestramento e inferenza dei modelli IA, e rappresentano oggi una delle maggiori barriere d'accesso alla competitività tecnologica. L'analisi della Bank for

International Settlements sottolinea come il mercato dei chip sia altamente concentrato e soggetto a tensioni geopolitiche[6].

Il *layer* ***cloud computing*** fornisce l'infrastruttura di calcolo necessaria per addestrare e distribuire i modelli. Un mercato dominato da pochi operatori globali, tra cui Amazon Web Services (AWS), Microsoft Azure, Google Cloud e Alibaba Cloud. Il controllo delle infrastrutture *cloud* è un nodo strategico, poiché determina chi può effettivamente operare su larga scala con modelli avanzati[7].

Il *layer* dei ***training data*** è costituito dai dataset utilizzati per addestrare i modelli. La qualità, l'ampiezza e la provenienza dei dati condizionano profondamente le *performance* dell'IA. La disponibilità di dati pubblici, commerciali o sintetici è oggi un elemento discriminante. Le barriere normative (come il GDPR) e il controllo dei flussi di dati rappresentano variabili critiche nella definizione della sovranità digitale[8].

Il *layer* dei ***foundation models*** riguarda i modelli pre-addestrati di grandi dimensioni, come GPT di OpenAI, LLaMA di Meta,

[6] *The AI supply chain*, BIS Papers No. 154 (by Leonardo Gambacorta and Vatsala Shreeti), March 2025, https://www.bis.org/publ/bppdf/bispap154.pdf.

[7] *The AI supply chain*, BIS Papers No. 154, op.cit.

[8] *The AI supply chain*, BIS Papers No. 154, op.cit.

Claude di Anthropic o Gemini di Google DeepMind. Questi modelli, una volta sviluppati, vengono adattati per compiti specifici attraverso il *fine-tuning*. La loro creazione richiede enormi risorse computazionali e finanziarie e sta concentrando il potere tecnologico in poche mani[9].

Infine, il *layer* delle **applicazioni IA** comprende le soluzioni verticali sviluppate a partire dai *foundation models* e adattate a specifici settori industriali, come sanità, finanza, pubblica amministrazione o manifattura. È a questo livello che l'intelligenza artificiale diventa un vero strumento trasformativo nei processi operativi e decisionali.

Ogni *layer* rappresenta un nodo critico di controllo economico e geopolitico. Il dominio statunitense sui primi due livelli (*hardware* e *cloud computing*) comporta una forte dipendenza strutturale per altri paesi. La Cina, consapevole di questa vulnerabilità, ha avviato piani di autosufficienza tecnologica che puntano a garantire il controllo su chip, dati e modelli. L'Unione Europea - pur in posizione subordinata nei livelli infrastrutturali - cerca di affermare la propria influenza tramite l'AI Act per quanto

[9] *The AI supply chain*, BIS Papers No. 154, op.cit.

concerne la regolazione di standard applicativi o, anche, attraverso altre iniziative come il più recente "AI Continent Action Plan" per puntare sull'innovazione e sul rafforzamento in termini tecnologici e infrastrutturali, sulla spinta agli investimenti e sul potenziamento delle competenze, al fine di consentire la realizzazione di un'IA affidabile per un'Europa più competitiva. **Comprendere questa stratificazione consente di leggere in chiave geopolitica e industriale le dinamiche di potere, che si stanno giocando attorno all'intelligenza artificiale**.

Nonostante i progressi tecnologici registrati negli ultimi decenni, la crescita della produttività a livello globale ha invece mostrato una tendenza stagnante, un fenomeno comunemente definito "*paradosso della produttività*"[10]. Tale apparente disallineamento tra l'innovazione tecnologica e il miglioramento dell'efficienza economica ha sollevato interrogativi sulla capacità delle nuove tecnologie di generare incrementi tangibili nella produzione e nella competitività in tempi rapidi. Almeno parte del potenziale inespresso dalla tecnologia è anzitutto

[10] S. DA EMPOLI, "Sbloccare produttività e crescita in Italia: l'IA è la risposta?", *AgendaDigitale.eu*, 11 giugno 2024, https://www.agendadigitale.eu/mercati-digitali/sbloccare-produttivita-e-crescita-in-italia-lia-e-la-risposta/.

dovuto, al momento, alla scarsa diffusione di competenze digitali, fenomeno che riguarda anche l'Italia[11].

In questo contesto, l'intelligenza artificiale è sempre più considerata la tecnologia a uso generale in grado di invertire questa traiettoria, poiché non si limita all'automazione dei processi fisici - come avvenuto nelle precedenti rivoluzioni industriali - ma interviene in maniera decisiva nell'automazione delle attività cognitive. Attraverso l'elaborazione avanzata dei dati, l'ottimizzazione delle decisioni e la riduzione dell'incertezza informativa, l'IA promette di sbloccare nuove frontiere di produttività, ridefinendo il rapporto tra capitale e lavoro e favorendo una più efficiente allocazione delle risorse nei sistemi economici.

L'adozione dell'intelligenza artificiale nell'economia globale segue due traiettorie distinte, riflettendo le differenze strutturali tra modelli economici e strategie nazionali. Da un lato, nelle economie occidentali - in particolare negli Stati Uniti e nei principali paesi europei - lo sviluppo dell'IA è in prevalenza

[11] A. FINOCCHIARO, "Digital Decade Report 2024: UE e Italia a confronto sul digitale", *FPA Digital 360*, 1 agosto 2024, https://www.forumpa.it/pa-digitale/digital-decade-report-2024-ue-e-italia-a-confronto-sul-digitale/.

trainato dalle imprese private. Il settore tecnologico, guidato da grandi attori come Microsoft, Google, OpenAI e NVIDIA, ha investito ingenti capitali nello sviluppo di modelli avanzati, finanziati da mercati privati e *venture capital*[12]. La competizione tra queste aziende ha accelerato l'innovazione, portando a un proliferare di modelli sia *open-source* che *closed-source*, ciascuno con implicazioni strategiche ed economiche differenti.

Un esempio significativo di approccio *open-source* è rappresentato dalla strategia adottata da Meta, che ha reso disponibili i modelli Llama 2 per la ricerca e lo sviluppo, favorendo una maggiore diffusione dell'IA e incoraggiando l'innovazione collaborativa[13]. Tale scelta è in netto contrasto con l'approccio *closed-source* di OpenAI, che - pur avendo inizialmente promosso un modello aperto - ha progressivamente limitato l'accesso ai suoi algoritmi avanzati per mantenere un vantaggio competitivo e proteggere gli interessi commerciali.

[12] G. MASI, "AI Index 2025: dov'è arrivata l'intelligenza artificiale, ecco la mappa totale", *AgendaDigitale.eu*, 10 aprile 2025, https://www.agendadigitale.eu/industry-4-0/ai-index-2025-ecco-tracciati-tutti-i-progressi-dellintelligenza-artificiale/.

[13] Il modello è reperibile sul sito di META all'indirizzo: https://www.llama.com/llama2/.

Dall'altro lato, in paesi come la Cina, l'adozione dell'intelligenza artificiale è fortemente sostenuta e diretta dallo Stato, con ingenti investimenti pubblici volti a consolidare la supremazia tecnologica nazionale. Il governo cinese ha varato piani strategici a lungo termine, come il programma "New Generation AI Development Plan", volto a rendere la Cina leader mondiale nell'IA entro il 2030[14]. In questo contesto, aziende come Baidu, Alibaba e Tencent collaborano strettamente con le autorità governative per sviluppare e applicare modelli di IA in ambiti strategici, come la sorveglianza, la difesa e la *governance* digitale[15]. L'approccio cinese predilige un modello *closed-source* per i dati, in cui l'accesso alle informazioni e agli algoritmi è rigidamente controllato dallo Stato per motivi di sicurezza nazionale e competitività economica. Sebbene alcune tecnologie e *framework* IA siano rilasciati in *open-source* (per esempio

[14] G. IUVINALE, N. IUVINALE, "IA, così Pechino punta al primato in tutti i campi: i rischi per l'Occidente, *AgendaDigitale.eu*, 4 luglio 2023, https://www.agendadigitale.eu/mercati-digitali/ia-cosi-pechino-punta-al-primato-in-tutti-i-campi-i-rischi-per-loccidente/.

[15] B. LARSEN, "Drafting China's National AI Team for Governance, *Digichina Stanford*, November 18, 2019, https://digichina.stanford.edu/work/drafting-chinas-national-ai-team-for-governance/; "The AI revolution: How China's government and private sector are pushing AI development forward", *CKGSB Knowledge*, April 28, 2025, https://english.ckgsb.edu.cn/knowledge/article/china-ai-leap-deepseek-govt-deive-ai-development/.

DeepSeek[16]), il controllo statale sui dataset, soprattutto quelli sensibili o strategici, rimane una priorità assoluta. Questo modello limita la circolazione dei dati al di fuori del paese e impone restrizioni severe sul loro utilizzo, favorendo lo sviluppo interno e riducendo la dipendenza da fornitori stranieri.

Questa biforcazione tra modelli *corporate-driven* e *government-driven* ha implicazioni profonde sulla regolamentazione dell'intelligenza artificiale e sulla distribuzione del potere economico a livello globale. Come di recente osservato[17], l'affermazione di attori privati nell'IA non implica soltanto una trasformazione tecnologica, ma ridefinisce i rapporti di potere tra individui, imprese e istituzioni democratiche, rendendo urgente una nuova infrastruttura pubblica per la ricerca e la *governance* dell'intelligenza artificiale.

Nei paesi occidentali, a partire dagli Stati Uniti, il dibattito si concentra su normative che bilancino innovazione e protezione dei dati, è il caso anche del regolamento europeo AI Act, volto a

[16] C. CRESCENZI, "DeepSeek, l'assistente AI cinese supera ChatGpt nell'App Store statunitense", *Wired*, 27 gennaio 2025, https://www.wired.it/article/deepseek-app-store/.

[17] G. AMATO & P. CONTUCCI, "La ricerca in intelligenza artificiale tra libertà e potere", *Rivista Il Mulino Online*, marzo 2025.

garantire trasparenza e sicurezza nell'uso dell'intelligenza artificiale. In paesi con un controllo statale più marcato, come la Cina, le politiche pubbliche sono più orientate alla massimizzazione dell'efficacia dell'IA come strumento di potere economico e geopolitico, con regolamenti che favoriscono lo sviluppo interno e meno la cooperazione esterna.

È una dicotomia che non solo influisce sulla velocità e sulla direzione dell'innovazione tecnologica, ma definisce anche i futuri equilibri di potere nell'economia globale, delineando scenari in cui l'accesso e il controllo dell'intelligenza artificiale diventeranno fattori chiave nella competizione internazionale.

Si sta tuttavia delineando una possibile terza via, incarnata dall'Unione Europea, in cui lo Stato non esercita un controllo diretto sui modelli e sui dati, ma assume un ruolo attivo nel promuovere, facilitare e sostenere lo sviluppo dell'intelligenza artificiale. L'intervento pubblico si concretizza attraverso investimenti strategici, programmi di ricerca, incentivi all'innovazione e quadri regolatori orientati alla tutela dei diritti fondamentali. **Questa via europea mira a conciliare l'autonomia tecnologica con i principi democratici**, evitando sia l'egemonia

delle *big tech* private, sia la centralizzazione autoritaria tipica di altri contesti geopolitici.

I.1.3 Rischi e opportunità dell'integrazione massiva dell'IA

L'IA nel contesto globale - cui è necessario riferirsi costantemente per comprenderne con chiarezza i rischi e le opportunità - presenta sistematicamente un volto ancipite. Considerate le sue enormi potenzialità in ogni campo dell'esperienza e dello scibile umani, per un osservatore attento emerge una polarizzazione estrema tra un orientamento indirizzato al miglioramento generale della società, dell'economia, della scienza, dei processi politici e decisionali e una loro possibile involuzione.

In questa sezione verrà posta l'attenzione su alcuni elementi relativi al mercato del lavoro che, nel breve periodo, appare come il primo a essere investito dalla rivoluzione digitale in corso.

Il rapido sviluppo dell'IA è destinato, come ogni innovazione tecnologica, a produrre un sostanziale cambiamento nel mercato

del lavoro, accentuando la perdita di potere contrattuale da parte dei lavoratori di alcuni settori. Già all'inizio del XXI secolo, la capacità dei capitali di spostarsi rapidamente - grazie alla nuova economia globalizzata, in quelle zone del pianeta dove le condizioni salariali e legislative erano molto più vantaggiose - aveva innescato una ridefinizione sostanziale delle dinamiche relative alla dialettica tra capitale e lavoro tipica del XX secolo. A ciò si aggiunge, adesso, la possibilità sempre più concreta di sostituire, quasi del tutto, la presenza di operatori umani all'interno della filiera industriale, in relazione a mansioni ripetitive, con una combinazione tra IA e robotica, destinata a migliorare l'efficienza produttiva.

Molti sono gli studi che insistono rispetto al miglioramento atteso della produttività mediato da intelligenza artificiale e robotica, alcuni dei quali, però, ridimensionano, anche di molto, le aspettative sull'impatto dell'intelligenza artificiale sulla produttività e sul lavoro. Esemplificativo, in questo senso, il lavoro dello scienziato statunitense Daron Acemoğlu, di recente insignito del premio Nobel per l'economia, che stima che

l'adozione dell'IA porterà a un aumento del PIL tra l'1,1% e l'1,6% nei prossimi dieci anni, con un incremento annuo della produttività di circa 0,05%[18]. Questa stima è in parte giustificata da quella che lo scienziato reputa essere un'aspettativa tradita dell'intelligenza artificiale generativa, che, nonostante le premesse iniziali non ha ancora prodotto applicazioni industriali rivoluzionarie e che trasformino significativamente il modo in cui si lavora[19].

Oltre che sulla produttività in generale, è opportuno evidenziare che **non esiste, ad oggi, un consenso nella letteratura economica e occupazionale sull'effetto reale che l'intelligenza artificiale avrà sul lavoro** in particolare. Le previsioni sull'automazione radicale o sulla cancellazione netta di intere mansioni non sono uniformemente validate da dati empirici.

Più che una scomparsa del lavoro, si delinea uno scenario di transizione, in cui il problema centrale è rappresentato da un

[18] D. ACEMOĞLU, "The simple macroeconomics of AI", *Economic Policy*, Vol. 40, Issue 121, January 2025, pp. 13–58, https://academic.oup.com/economicpolicy/article/40/121/13/7728473?login=false.

[19] P. DIZIKES, "What do we know about the economics of AI? Nobel laureate Daron Acemoglu has long studied technology-driven growth", *MIT News*, December 6, 2024, https://economics.mit.edu/news/daron-acemoglu-what-do-we-know-about-economics-ai.

mismatch temporaneo tra le competenze richieste dai nuovi processi tecnologici e quelle attualmente disponibili nel mercato del lavoro. Questo aspetto è ampiamente trattato nel *Future of Jobs Report 2023* del World Economic Forum, che stima che il 44% delle competenze dei lavoratori cambierà nei prossimi cinque anni, sottolineando come il disallineamento tra domanda e offerta di skill rischi di produrre disoccupazione frizionale e nuove vulnerabilità sociali[20].

Una posizione analoga è espressa nel rapporto dell'OCSE *Artificial Intelligence, Job Quality and Inclusiveness* (2023)[21], dove si afferma che l'adozione dell'IA comporterà un riassetto profondo delle mansioni, più che una distruzione netta dei posti di lavoro. Tuttavia, in assenza di politiche coordinate tra sistema educativo, imprese e istituzioni pubbliche, il rischio è quello di amplificare le disuguaglianze e compromettere la qualità dell'inclusione nel mercato del lavoro.

[20] WORLD ECONOMIC FORUM, *Future of jobs report 2023*, Davos, May 2023, https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf.

[21] OECD, *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, 2023, cfr. https://www.oecd.org/en/publications/oecd-employment-outlook-2023_08785bba-en/full-report/artificial-intelligence-job-quality-and-inclusiveness_a713d0ad.html.

Sull'impatto dell'interazione tra lavoratori e intelligenza artificiale è opportuno citare un ulteriore lavoro recente. Un recente studio condotto da ricercatori di Harvard Business School, Wharton School e Procter & Gamble ha esplorato il concetto di *cybernetic teammates*[22], ovvero sistemi di intelligenza artificiale progettati per integrarsi attivamente in contesti collaborativi, potenziando le dinamiche di gruppo e la condivisione delle competenze.

L'esperimento ha coinvolto 776 professionisti impegnati in sfide reali di innovazione di prodotto. I partecipanti sono stati assegnati casualmente a lavorare con o senza l'ausilio dell'IA, sia individualmente che in squadra. I risultati hanno mostrato che gli individui che utilizzavano l'IA raggiungevano prestazioni comparabili a quelle delle squadre senza IA, evidenziando come l'IA possa replicare molti dei benefici della collaborazione umana. Inoltre, l'IA ha contribuito a superare i silos funzionali: i

[22] DELL'ACQUA, F., AYOUBI, C., LIFSHITZ-ASSAF, H., SADUN, R., MOLLICK, E. R., MOLLICK, L., HAN, Y., GOLDMAN, J., NAIR, H., TAUB, S., & LAKHANI, K. R. (2025), "The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise", *Harvard Business School Working Paper* No. 25-043, Disponibile su SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5188231.

professionisti, indipendentemente dal loro *background* tecnico o commerciale, hanno prodotto soluzioni più bilanciate quando coadiuvati dall'IA.

Questi risultati suggeriscono che l'adozione dell'IA su larga scala nel lavoro cognitivo non solo migliora le prestazioni, ma ridefinisce anche il modo in cui la competenza e la connettività sociale si manifestano all'interno delle squadre, spingendo le organizzazioni a ripensare la struttura stessa del lavoro collaborativo.

È dunque legittimo chiedersi se questa tecnologia renderà nel tempo sempre più superfluo quanto resta del lavoro a basso tasso di specializzazione e manualità e se - più in generale - avrà un impatto sulle macro-dinamiche del mercato del lavoro, sulla formazione, sull'acquisizione e la distribuzione delle competenze. In questo contesto, la comprensione politica del fenomeno diventa essenziale: senza una lettura sistemica, il ritardo nell'adattamento delle competenze rischia di amplificare tensioni sociali e diseguaglianze.

A sottolineare questa urgenza è anche il rapporto *Artificial Intelligence and the future of Work*, pubblicato dall'Eurobarometer, che rileva anch'esso l'opportunità di adottare una *governance* multilivello capace di legare visione strategica, investimenti educativi e strumenti redistributivi[23].

A tal riguardo, tuttavia, si può realisticamente prevedere la creazione di nuove figure professionali e la rivalutazione sostanziale di tutte quelle attività artigianali fondate su una manualità difficilmente riproducibile, nel prossimo futuro, a livello robotico. Quest'ultimo genere di occupazioni conoscerà probabilmente un nuovo e rinnovato interesse da parte del mercato e un suo relativo apprezzamento sul piano sociale ed economico.

Per quanto concerne, invece, la creazione di nuove professionalità si può ipotizzare che esse saranno centrate sul controllo del comportamento dell'IA da parte dell'operatore umano e sulla sua sostanziale ibridazione con esso. In tal senso è di importanza

[23] *Artificial Intelligence and the future of work*, Special Eurobarometer Report 554, April-May 2024, https://europa.eu/eurobarometer/api/deliverable/download/file?deliverableId=96284; Si veda anche questo *report* che comprende gli Stati Uniti: *The Impact of Artificial Intelligence on the Future of Workforces in the EU and the US*, 5 December 2022, https://digital-strategy.ec.europa.eu/en/library/impact-artificial-intelligence-future-workforces-eu-and-us

fondamentale la riqualificazione di lavoratori e professionisti attraverso programmi di formazione a loro specificamente dedicati, prevedendo un reddito di transizione opportunamente calibrato e promuovendo, grazie ad appositi incentivi, lo sviluppo del lavoro integrato uomo-macchina in grado di generare nuove opportunità professionali.

A tutto ciò si può aggiungere: sul piano legislativo - allo scopo di mitigare l'impatto trasformativo delle nuove tecnologie - l'obbligo per le aziende di garantire ai lavoratori l'accesso a percorsi di aggiornamento professionale; a livello operativo, lo sviluppo di nuove professionalità, investendo nei settori emergenti dell'etica dell'IA, della scienza dei dati e della manutenzione dei sistemi informatici (*software* e *hardware*).

Le competenze richieste nello specifico sono di ordine squisitamente tecnico, di tipo trasversale e di alfabetizzazione digitale. Nel primo caso si tratta di programmazione, *machine learning*, gestione dei dati e *cybersecurity*; nel secondo di *problem solving*, creatività e adattamento mentale; nel terzo di comprensione dei limiti e delle potenzialità dell'IA in relazione alle sue applicazioni e, di conseguenza, dell'acquisizione di una

conoscenza approfondita del dominio applicativo di riferimento da parte dell'operatore umano.

Allo scopo di rispondere alle nuove esigenze dal punto di vista della formazione scolastica e professionale si ritiene necessario introdurre, poi - anche a livello accademico e dell'istruzione secondaria superiore - alcuni percorsi seminariali focalizzati sul corretto approccio cognitivo all'IA e alle soluzioni informatiche d'avanguardia[24]. Ciò allo scopo di evitare che un loro uso massivo e sconsiderato produca in tempi rapidi un deterioramento e una sostanziale amputazione delle facoltà mentali legate al pensiero convergente (logico-analitico).

Tali percorsi dovrebbero essere parte di una più ampia strategia formativa, trasversale a tutti i livelli di cultura e di età, volta a

[24] Su questo tema si segnalano questi tre studi: WORLD ECONOMIC FORUM, *Shaping the Future of Learning: The Role of AI in Education 4.0*, Insight Report, April 2024, https://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Learning_2024.pdf; si veda anche questa analisi che parte da un sondaggio su un campione di quindici paesi, indicando alcune competenze che i governi potrebbero voler prioritizzare: M. DONDI, J. KLIER, F., J. SCHUBERT, *Defining the skills citizens will need in the future world of work*, McKinsey & Company, June 2021, https://www.mckinsey.com/~/media/mckinsey/industries/public%20and%20social%20sector/our%20insights/defining%20the%20skills%20citizens%20will%20need%20in%20the%20future%20world%20of%20work/defining-the-skills-citizens-will-need-in-the-future-of-work-final.pdf; L. CAO & C. DEDE, *Navigating A World of Generative AI: Suggestions for Educators*, Next Level Lab at Harvard Graduate School of Education, President and Fellows of Harvard College: Cambridge, MA, https://bpb-us-e1.wpmucdn.com/websites.harvard.edu/dist/a/108/files/2023/08/Cao_Dede_final_8.4.23.pdf.

modificare la mentalità (*mindset*) e la cultura della società nel suo complesso, realizzata anche attraverso la collaborazione tra aziende e istituzioni educative. Andrà anche valutata l'opportunità, per quanto concerne la formazione tecnico-professionale, di creare percorsi coerenti con le dinamiche tipiche del mercato del lavoro.

In tal senso, **si rende urgente una riforma dei metodi scolastici, un'adeguata formazione degli insegnanti e un avvicinamento dei bambini, fin dalle scuole elementari, a un uso consapevole, intelligente ed etico dell'intelligenza artificiale**. Un esempio interessante in questa direzione proviene dalla Cina, dove a partire dal 2025 è previsto l'inserimento obbligatorio dell'insegnamento dell'IA già nelle scuole primarie, con programmi didattici che includono nozioni pratiche e riflessioni etiche fin dall'infanzia[25] .

È importante sottolineare che l'intelligenza artificiale generativa rappresenta uno strumento complesso: se da un lato può essere estremamente utile per facilitare l'apprendimento, dall'altro può

[25] L. CHONG MING, "China's capital city is making AI education mandatory, even for elementary schoolers", *Business Insider*, March 10, 2025, https://www.businessinsider.com/china-beijing-ai-education-mandatory-classrooms-elementary-schoolers-2025-3.

anche costituire un limite se utilizzata in modo acritico. A tal riguardo, la teoria del "*Sistema 0*", proposta da Massimo Chiriatti e Giuseppe Riva, descrive l'IA come una sorta di estensione cognitiva esterna che elabora informazioni senza comprenderle, necessitando quindi dell'intervento umano per attribuire significato ai dati processati. Tale concetto evidenzia l'importanza di mantenere un controllo critico sull'uso dell'IA, evitando che essa sostituisca completamente il pensiero umano[26]. Il che è un'ulteriore ragione per voler investire su un quadro di norme etiche imperniate sulla responsabilità attiva del soggetto che crea le applicazioni, essendo queste sempre più vicine ad un nuovo "strato cognitivo", il quale non solo orienta le scelte, ma rischia di sostituire la volontà e il pensiero autonomo con una accettazione passiva delle proposte fornite dall'IA e fondate su abitudini e preferenze accumulate nel tempo.[27]

[26] CHIRIATTI, M., GANAPINI, M., PANAI, E., UBIALI, M., & RIVA, G. (2024), "The case for human–AI interaction as system 0 thinking", *Nature Human Behaviour*, 8(10), 2024, pp. 1829-1830; cfr. anche: C. GALLETTI, "Sistema 0, l'intelligenza artificiale sta già cambiando il cervello umano: cosa ha scoperto una ricerca italiana", *Log in - Corriere della Sera*, 26 ottobre 2024, https://www.corriere.it/tecnologia/24_ottobre_26/sistema-0-l-intelligenza-artificiale-sta-gia-cambiando-il-cervello-umano-cosa-ha-scoperto-una-ricerca-italiana-dcb525d8-7b4e-42f6-8ed2-6d9f4f1a5xlk.shtml.

[27] M. CHIRIATTI, G. Riva, "Sistema 0: l'intelligenza artificiale che sta trasformando il nostro modo di pensare", *Pandora Rivista*, 22 aprile 2025, https://www.pandorarivista.it/articoli/sistema-0-l-intelligenza-artificiale-che-sta-trasformando-il-nostro-modo-di-pensare/.

Se si considera la rapidità con cui evolvono i contenuti e le competenze richieste nel mondo del lavoro, è forse opportuno ripensare la scuola non più come un luogo dove si assorbono passivamente informazioni, ma come un ambiente in cui si impara ad imparare e ad affrontare in maniera seria e precisa problemi complessi. In questo contesto, la celebre frase attribuita a Plutarco: "*i giovani sono fuochi da accendere e non vasi da riempire*", assume un significato ancora più profondo, sottolineando l'importanza di stimolare la curiosità e il pensiero critico degli studenti.

A livello internazionale, sono stati sviluppati diversi modelli di scuola moderna nell'era dell'IA. Ad esempio, il progetto "*Classe AI*", promosso da Fondazione Vodafone, Fondazione Golinelli e Stand Up City, offre soluzioni digitali per docenti, studenti e istituti scolastici, con l'obiettivo di integrare l'IA nella didattica in modo efficace e responsabile[28]. Inoltre, iniziative come *TeachAI* forniscono *toolkit* pratici per le scuole, aiutando gli

[28] Progetto CLASSE AI - didattica innovativa con l'intelligenza artificiale, cfr. https://www.classeai.it/.

educatori a comprendere e adottare l'IA nell'insegnamento, promuovendo un approccio etico e centrato sullo studente[29].

Oltre all'ampliamento dell'offerta formativa in senso tecnico o specialistico, è fondamentale adottare un nuovo paradigma culturale, che promuova l'apprendimento continuo (*lifelong learning*) e trasformi le organizzazioni anche in "*learning companies*". Questo approccio è stato efficacemente adottato a Singapore attraverso il programma nazionale *SkillsFuture*, che offre a tutti i cittadini opportunità di sviluppo delle competenze lungo l'intero arco della vita, indipendentemente dal punto di partenza. Il programma include crediti formativi, corsi approvati e iniziative mirate a diverse fasce d'età e categorie professionali[30].

In Europa, la Commissione Europea ha delineato strategie per l'apprendimento permanente, sottolineando l'importanza di competenze chiave come alfabetizzazione, multilinguismo,

[29] L. BENUSSI, "Intelligenza Artificiale, scuola e pensiero creativo: stiamo preparando i giovani alle professioni del futuro?", *Forum PA*, 8 febbraio 2024, https://www.forumpa.it/temi-verticali/scuola-istruzione-ricerca/intelligenza-artificiale-scuola-e-pensiero-creativo-stiamo-preparando-i-giovani-alle-professioni-del-futuro/.

[30] Per la piattaforma *Skillsfuture* v. https://www.skillsfuture.gov.sg/aboutskillsfuture.

competenze digitali e imprenditoriali. Tali competenze sono considerate essenziali per la realizzazione personale, l'occupabilità e la cittadinanza attiva[31].

In Italia, il Fondo Nuove Competenze[32] rappresenta un tentativo di promuovere la formazione continua, consentendo alle imprese di rimodulare l'orario di lavoro per attività formative. Tuttavia, il programma presenta alcune limitazioni, tra cui la complessità burocratica e la necessità di una maggiore integrazione con le strategie aziendali di lungo termine[33]. In questo contesto, è fondamentale anche il ruolo degli enti bilaterali, che offrono servizi di formazione continua e aiutano l'incontro tra domanda e offerta di lavoro. Costituiti da associazioni datoriali e sindacali, questi organismi promuovono iniziative in materia di formazione e riqualificazione

[31] EUROPEAN COMMISSION, European Education Area, https://education.ec.europa.eu/it/education-levels/school-education/basic-skills.

[32] Sul sito del Ministero del Lavoro e delle Politiche sociali, si veda il focus sul Fondo nuove competenze, https://www.lavoro.gov.it/temi-e-priorita/orientamento-e-formazione/focus/fondi-alle-imprese-la-formazione-continua/pagine-0.

[33] PRIMOPIANO, Fondo Nuove Competenze 2024 - 2025: guida per aziende, https://corsi.primopiano.it/fondo-nuove-competenze/.

professionale, contribuendo alla creazione di un ecosistema di apprendimento collaborativo.

In tale scenario, lo Stato può assumere il ruolo di facilitatore, promuovendo politiche pubbliche volte a incentivare l'apprendimento continuo e a sostenere la trasformazione delle organizzazioni anche in "*learning companies*". Ciò include la creazione di infrastrutture per l'accesso equo all'apprendimento, la definizione di quadri normativi a favore di una formazione permanente e della promozione di una cultura dell'apprendimento come valore sociale. In alcuni paesi, si va oltre l'idea di formazione come attività individuale o aziendale per adottare modelli sistemici di apprendimento. In Finlandia, ad esempio, il sistema educativo è costruito come un vero e proprio ecosistema di apprendimento basato su flessibilità, responsabilità locale e collaborazione tra attori pubblici, scuole e imprese. Programmi come *Yrityskylä* coinvolgono direttamente gli studenti in esperienze pratiche, introducendoli a dinamiche economiche e lavorative reali[34].

[34] R. MILNE, "Finland fuels childrens future", *The Financial Times*, January 6 2025, https://www.ft.com/content/26c56174-76ab-493b-9770-6d1ed4996505.

Anche nei Paesi Bassi si osserva una crescente attenzione verso l'integrazione dell'apprendimento per lo sviluppo sostenibile in un quadro coerente e collaborativo, che coinvolga scuole, comunità e stakeholder territoriali[35].

A livello internazionale, organizzazioni come l'OCSE[36], l'UNESCO[37] e la piattaforma europea EPALE[38] promuovono il concetto di *learning ecosystems* come reti locali di apprendimento interconnesse, capaci di valorizzare ogni contesto - dalla scuola al lavoro, dal digitale al volontariato - come spazi formativi. In questa visione, l'apprendimento non è confinato in un'aula o in un'età della vita, ma diventa una dimensione costante e distribuita, da sostenere con infrastrutture, politiche e alleanze territoriali.

[35] EUROPEAN COMMISSION, Education and Training Monitor 2024, scheda paese Olanda, https://op.europa.eu/webpub/eac/education-and-training-monitor/en/country-reports/netherlands.html.

[36] OECD, *Country Digital Education Ecosystems and Governance: A Companion to Digital Education Outlook 2023*, OECD Publishing, Paris, 2023, https://www.oecd.org/en/publications/country-digital-education-ecosystems-and-governance_906134d4-en.html.

[37] UNESCO, "Lifelong learning ecosystems must be inclusive of all learners", press release, 5 July 2023, https://www.uil.unesco.org/en/articles/lifelong-learning-ecosystems-must-be-inclusive-all-learners.

[38] S. GEVAERTS, "Learning appetite nourished by an ecosystem", *EPALE Blog*, 7 June 2022, https://epale.ec.europa.eu/en/blog/learning-appetite-nourished-ecosystem.

Dal punto di vista degli strumenti tecnici a sostegno, si ritiene opportuno potenziare le piattaforme di apprendimento *on-line*, attraverso la diffusione di corsi MOOC (*Massive Open Online Course*) per l'aggiornamento delle competenze digitali e professionali.

Appare quindi abbastanza probabile come nel prossimo futuro vi sarà inevitabilmente una polarizzazione all'interno del mercato del lavoro tra professionisti altamente qualificati che grazie alle loro competenze - opportunamente potenziate dall'uso dell'IA e delle nuove tecnologie - vedranno crescere opportunità e retribuzioni e lavoratori a bassa qualificazione, che invece subiranno conseguenze diametralmente opposte. Tutto ciò porterà a inevitabili conseguenze di tipo socioeconomico che incrementeranno la sperequazione sociale e, pertanto, sarà necessario predisporre diverse linee di intervento: per un verso, attraverso l'erogazione di sussidi pubblici e, per un altro, sul piano della ricollocazione all'interno del mercato di quei soggetti che ne rimarranno esclusi.

Si possono a tal riguardo immaginare alcune strategie operative come incentivi alla formazione, creazione di nuovi ruoli nei

settori dei servizi digitali e di sostegno all'IA, politiche di inclusione digitale e la creazione di un quadro normativo in grado di limitare, per quanto possibile, pratiche discriminatorie di vario genere, concernenti l'uso e l'accesso all'IA, l'assunzione di personale e la sua gestione.

Per quanto concerne l'orizzonte regolatorio, sembra necessario un quadro normativo condiviso all'interno dello spazio politico e culturale della civiltà occidentale, che sia però dotato della necessaria flessibilità coerente con la rapida evoluzione delle tecnologie digitali. Attualmente si assiste, invece, al proliferare di atti regolatori tra loro assai eterogenei, tra cui spicca l'AI Act approvato dall'UE.

Quest'ultimo intervento in particolare - al di là del merito dei principi ivi contenuti - deve essere affiancato dal potenziamento della *leadership* europea nella realizzazione di sistemi di IA, considerando che, al momento, sono gli Stati Uniti e la Cina ad essere i principali player nel settore. **Solo a condizione di sviluppare un'IA efficace e competitiva, prodotta da aziende**

che rispondono del loro operato all'UE, sarà possibile far rispettare tale regolamentazione e renderla realmente efficace.

Rimane in ogni caso di difficile soluzione la questione connessa con le sfide etiche, politiche e sistemiche che sono inevitabilmente innescate dall'odierno rapido sviluppo tecnologico. Si tratta di un problema di fondo relativo all'allineamento tra obiettivi umani, conservazione della formula politica liberal democratica, necessità per le aziende di produrre profitti apprezzabili in un quadro di competizione sfrenata per la supremazia tecnologica e comportamento a tratti imprevedibile dell'IA stessa.

Tutti gli elementi considerati rischiano di non trovare una composizione equilibrata e la loro interazione potrebbe condurre a esiti imprevedibili, soprattutto in mancanza di un quadro etico-normativo largamente condiviso. A tal proposito sarebbe opportuno favorire la cooperazione internazionale a livello governativo, accademico e sociale.

I.2. FOCUS SUI SETTORI CHIAVE DI ADOZIONE DELL'IA

In questa sezione verranno approfondite le applicazioni legate all'utilizzo dell'intelligenza artificiale in alcuni settori strategici o ad alto potenziale, nei quali l'impatto della tecnologia è già osservabile o potrebbe determinare trasformazioni rilevanti nel breve-medio periodo. In tal senso, l'adozione dell'IA non si configura come un'evoluzione marginale, ma come una leva per la riconfigurazione di modelli di produzione, erogazione di servizi e interazione tra attori pubblici e privati. Saranno analizzati, con taglio settoriale, i principali ambiti in cui l'intelligenza artificiale sta abilitando nuove efficienze, nuove logiche di funzionamento e nuove criticità, con un'attenzione particolare al contesto italiano ed europeo.

I.2.1 SANITÀ E MEDICINA: SCOPERTA DI NUOVE MOLECOLE E MEDICINA DI PRECISIONE

L'intelligenza artificiale sta progressivamente trasformando il settore sanitario, non solo migliorando la diagnosi e la ricerca di nuove terapie, ma soprattutto rafforzando la capacità dei sistemi sanitari di agire in chiave preventiva, attraverso l'analisi

predittiva di grandi quantità di dati. In uno scenario che prevede la sostenibilità messa alla prova dall'aumento delle cronicità, l'invecchiamento della popolazione e la pressione sui costi, l'IA si configura come una leva chiave per passare da un modello reattivo a uno proattivo e predittivo.

Tra i principali contributi dell'IA alla medicina del futuro vi è infatti la possibilità di anticipare l'evoluzione clinica dei pazienti. Attraverso algoritmi di *machine learning* applicati a dati genetici, comportamentali, clinici ed esogeni (come quelli ambientali e socioeconomici), è possibile stimare con elevata accuratezza il rischio individuale di sviluppare patologie cardiovascolari, metaboliche, neurodegenerative o tumorali.

Ad esempio, diversi studi scientifici recenti[39] hanno mostrato come un modello di IA multimodale per prevedere casi di

[39] SILCOX, C., ZIMLICHMANN, E., HUBER, K. *ET AL.*, "The potential for artificial intelligence to transform healthcare: perspectives from international health leaders", *npj Digit. Med*. 7, 88 (2024), https://doi.org/10.1038/s41746-024-01097-6. Sul Diabete si veda: ARSLAN, A. K., YAGIN, F. *ET AL.*, "Enhancing type 2 diabetes mellitus prediction by integrating …", *Frontiers in Endocrinology*, Vol. 15, 2024, https://www.frontiersin.org/journals/endocrinology/articles/10.3389/fendo.2024.1444282/full; S. BHUSHAN SINGH, A. SINGH, "Leveraging Deep Learning and Multi-Modal Data for Early Prediction and Personalized Management of Type 2 Diabetes", *IJFMR*, Vol. 6, Issue 4, July-August 2024, https://www.ijfmr.com/papers/2024/4/26096.pdf e S. ELLAHHAM, "Artificial Intelligence: The Future for Diabetes Care", *The American Journal of Medicine*, Vol. 133, Issue 8, 2020, pp. 895-900, https://www.sciencedirect.com/science/article/abs/pii/S0002934320303399.

diabete di tipo 2 - integrando note cliniche e dati di laboratorio - riesce a raggiungere livelli di accuratezza prossimi all'80%, anche in soggetti asintomatici.

Una capacità predittiva simile consente di attivare strategie preventive mirate - farmacologiche, nutrizionali o comportamentali - prima che si manifestino i sintomi. In questo modo, la prevenzione coadiuvata dall'IA non solo riduce i costi assistenziali futuri, ma migliora significativamente la qualità della vita e la prognosi dei pazienti.

L'ambito in cui tale trasformazione risulta particolarmente evidente è quello della medicina di precisione, che grazie all'IA diventa sempre più personalizzata. L'intelligenza artificiale consente infatti di superare l'approccio "*taglia unica*" (*one-fits-all*) alla cura, abilitando percorsi terapeutici disegnati sulla base del profilo individuale del paziente, che integra informazioni genomiche, molecolari, cliniche e ambientali.

Recenti studi su *Frontiers in Molecular Biosciences* (2025) e su istologie H&E hanno evidenziato l'efficacia di reti neurali

profonde nel riconoscere sottotipi tumorali del carcinoma prostatico, con implicazioni dirette per la scelta dei trattamenti[40].

L'elaborazione automatica di questi dati permette di identificare con maggiore precisione le terapie più efficaci e meglio tollerate per ciascun paziente, riducendo il *trial-and-error* terapeutico. A sostegno di questo approccio emergono anche soluzioni basate su modelli digitali del paziente - i cosiddetti "*digital twin*"- che consentono di simulare l'effetto di diverse opzioni terapeutiche e di ottimizzare la scelta clinica, come già sperimentato in contesti ospedalieri avanzati in Germania[41] e Singapore[42].

In questo contesto si inserisce la tendenza più ampia delle *healthness* e *longevity*, in cui la medicina di precisione dà un

[40] XIAOFENG HE *ET AL.*, "Integrative multi-omics analysis and machine learning refine global histone modification features in prostate cancer", *Frontiers in Molecular Biosciences*, Vol. 12 - 2025, https://www.frontiersin.org/journals/molecular-biosciences/articles/10.3389/fmolb.2025.1557843/full; E. ERAK, L. DEPAULA OLIVEIRA *ET AL.*, "Predicting Prostate Cancer Molecular Subtype With Deep Learning on Histopathologic Images", *Modern Pathology*, Vol. 36, Issue 10, Issue 10, 100247, 2023, https://www.sciencedirect.com/science/article/pii/S0893395223001527.

[41] "Fraunhofer Institutes present the first prototype for digital twins of patients", Research News, Press Release, November 2, 2021, https://www.fraunhofer.de/en/press/research-news/2021/november-2021/fraunhofer-institutes-present-the-first-prototype-for-digital-twins-of-patients.html.

[42] "Digital-twin tech for managing chronic kidney disease to be trialled in Singapore in early 2025", *The Straits Times*, 5 August, 2024, https://www.sgh.com.sg/news/patient-care/digital-twin-tech-for-managing-chronic-kidney-disease-to-be-trialled-in-singapore-in-early-2025.

orientamento che incide su un intero stile di vita orientato al benessere; tra le novità le "*med-cations*" vale a dire vacanze dedicate alla prevenzione e alla ricomposizione biologica[43].

Oltre alla personalizzazione e alla prevenzione, l'IA è già oggi impiegata in ambito sanitario per aumentare l'efficienza dei processi (per alcuni esempi di applicazioni v. anche *infra* § II.2). Algoritmi addestrati su immagini mediche sono di aiuto ai professionisti nella diagnosi radiologica[44]; sistemi di NLP automatizzano la produzione di referti e la gestione documentale[45]; modelli predittivi aiutano a pianificare l'allocazione delle risorse ospedaliere in base al rischio di ricovero o riammissione[46].

[43] M. GROSS, "Transformative 'med-cations' are the ultimate 2025 wellness trend", *New York Post*, June 12, 2025, https://nypost.com/2025/06/12/lifestyle/transformative-med-cations-are-the-ultimate-wellness-trend/. Si veda anche *infra* il paragrafo dedicato alle performance atletiche, § I.2.4.

[44] J. FRIEDLANDER SERRANO, "AI hasn't killed radiology, but it is changing it", *The Washington Post*, April 5, 2025, https://www.washingtonpost.com/health/2025/04/05/ai-machine-learning-radiology-software/.

[45] Si veda la scheda informativa a cura di foresee medical, *Natural language processing in healthcare*, https://www.foreseemed.com/natural-language-processing-in-healthcare.

[46] ROMERO-BRUFAU S., WYATT K.D. *ET AL.*, "Implementation of Artificial Intelligence-Based Clinical Decision Support to Reduce Hospital Readmissions at a Regional Hospital", *Appl Clin Inform.*, 2020 August; 11(4), pp. 570-577, https://pmc.ncbi.nlm.nih.gov/articles/PMC7467834/.

In alcuni casi, come nella cura domiciliare degli anziani, algoritmi predittivi già permettono di ridurre significativamente il numero di eventi avversi attraverso interventi tempestivi. È il caso del sistema adottato dal NHS in collaborazione con la società britannica Cera, che dichiara riduzioni fino al 70% delle ospedalizzazioni e del 20% delle cadute, con un risparmio stimato di oltre un milione di sterline al giorno[47]. Tuttavia, mentre queste applicazioni migliorano le performance operative, sono le capacità predittive e la medicina di precisione a rappresentare il vero potenziale trasformativo dell'IA nel modello sanitario nel suo complesso.

In Italia, l'adozione dell'IA in ambito medico è ancora contenuta. Secondo i dati ISTAT 2024[48], solo l'8,2% delle imprese italiane in generale utilizza sistemi basati sull'IA, una quota inferiore alla media europea. I limiti strutturali, come la frammentazione del

[47] NHS ENGLAND, "Nationwide roll out of artificial intelligence tool that predicts falls and viruses", 5 March 2025, https://www.england.nhs.uk/2025/03/nationwide-roll-out-of-artificial-intelligence-tool-that-predicts-falls-and-viruses/ vedi anche l'articolo "Healthcare start-up Cera wins unicorn status after raising $150m", *The Times*, January 13 2025, https://www.thetimes.com/business-money/companies/article/healthcare-start-up-cera-wins-unicorn-status-after-raising-150m-x6dczkmwc.

[48] ISTAT, "Un terzo delle grandi imprese utilizza tecnologie di Intelligenza Artificiale", *Statistiche Report*, 17 gennaio 2025, https://www.istat.it/wp-content/uploads/2025/01/Statreport_ICT2024-1.pdf.

sistema sanitario, la disomogeneità nell'adozione delle cartelle elettroniche e la scarsità di interoperabilità dei dati, rallentano la piena attuazione di soluzioni *data-driven* su cui invece l'Unione Europea ha incentrato la sua strategia attraverso il Data Governance Act e il concetto di *data altruism*.

Si aggiunge poi la necessità di rafforzare le competenze digitali tra i professionisti sanitari e di chiarire i requisiti normativi ed etici legati all'utilizzo dell'IA, in particolare su temi come la *privacy*, la trasparenza algoritmica e la responsabilità clinica. Il potenziale dell'IA nei sistemi sanitari europei sarà pienamente sbloccabile solo in presenza di una *governance* dei dati condivisa e di una solida cultura digitale tra i decisori pubblici. Iniziative europee e nazionali stanno iniziando ad affrontare questi nodi, ma resta ancora molta strada da fare per una reale trasformazione[49].

Affinché l'Italia possa cogliere appieno il potenziale dell'intelligenza artificiale in ambito medico, è necessario sviluppare un piano organico di investimento e *policy*. Tra le

[49] Per l'Italia si può vedere l'analisi *The impact of Artificial Intelligence in Italy from finance to healthcare*, Rome Business School Research, 10 July 2024, https://romebusinessschool.com/blog/the-impact-of-artificial-intelligence-in-italy-from-finance-to-healthcare/.

azioni prioritarie: la creazione di infrastrutture dati federate e sicure per la ricerca e la pratica clinica[50]; l'integrazione dell'IA nei percorsi di formazione medica e infermieristica; la promozione di progetti pilota su scala regionale per validare le soluzioni di medicina predittiva e personalizzata come anche l'adozione di standard nazionali per l'interoperabilità dei sistemi.

Oltre alla gestione del paziente, anche la *discovery* farmacologica è fortemente investita dall'intelligenza artificiale. Tradizionalmente, lo sviluppo di nuovi farmaci è caratterizzato da un processo lungo e oneroso. La sola fase di scoperta preclinica richiede generalmente tra i tre e i sei anni, con costi che possono ammontare a centinaia di milioni di euro, prima che una molecola candidata entri nelle sperimentazioni cliniche sull'uomo[51]. L'intelligenza artificiale sta radicalmente trasformando questa dinamica, riducendo tempi e costi in modo significativo.

[50] Sul tema si rimanda anche al Paper ASPEN INSTITUTE ITALIA coordinato da G.FINOCCHIARO, *La gestione del dato in sanità: nuove prospettive per il SSN*, Roma 24 ottobre 2024, (programma "Innovazione in Medicina") https://www.aspeninstitute.it/la-gestione-del-dato-in-sanita-nuove-prospettive-per-il-ssn/.

[51] M. CHUN, "How Artificial Intelligence is Revolutionizing Drug Discovery", The Petrie Flom Center, Harvard, March 20, 2023, https://petrieflom.law.harvard.edu/2023/03/20/how-artificial-intelligence-is-revolutionizing-drug-discovery/

Un esempio emblematico è rappresentato dal farmaco antifibrotico, sviluppato da *Insilico Medicine* nel 2022[52,53]: grazie all'uso dell'IA, il composto è passato dall'identificazione del bersaglio biologico alla fase I della sperimentazione clinica in meno di trenta mesi[54], un risultato straordinario rispetto ai tempi medi convenzionali. Tale progetto ha utilizzato modelli di *machine learning* per individuare un nuovo target biologico per la fibrosi polmonare idiopatica e generare una molecola farmaceutica ottimizzata. A titolo di confronto, un programma preclinico tradizionale per lo sviluppo di un farmaco simile può superare il miliardo di euro in costi capitalizzati. Ciò dimostra come gli strumenti fondati su IA siano in grado di analizzare immensi database chimici e biologici con una rapidità ineguagliabile rispetto ai ricercatori umani, individuando candidati farmaceutici promettenti in tempi record.

[52] F. REN, A. ALIPER *ET AL.*, "A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models", *Nature Biotechnology*, vol. 43, pp. 63–75, 2025, https://www.nature.com/articles/s41587-024-02143-0

[53] "Insilico Medicine's AI-driven drug Rentosertib receives official generic name", *News medical life sciences*, March 7, 2025, https://www.news-medical.net/news/20250307/Insilico-Medicines-AI-driven-drug-Rentosertib-receives-official-generic-name.aspx

[54] STOKES J.M., YANG K. *ET AL.*, "A Deep Learning Approach to Antibiotic Discovery", *Cell*. 2020, Feb 20; 180(4), pp. 688-702 https://insilico.com/phase1

Un ulteriore traguardo significativo è stata la scoperta di *Halicin*, un potente antibiotico individuato da un modello di IA sviluppato presso il Massachusetts Institute of Technology[55]. L'intelligenza artificiale ha esaminato oltre cento milioni di molecole in pochi giorni, un compito impraticabile con i metodi sperimentali tradizionali. Il modello ha identificato una molecola efficace contro diversi batteri resistenti agli antibiotici esistenti, nonostante tale composto non fosse mai stato considerato dai ricercatori per un uso antimicrobico.

Analogamente, la società britannica Exscientia[56] ha annunciato nel 2020 l'ingresso nei trial clinici della prima molecola farmaceutica progettata interamente da un'IA, destinata al trattamento del disturbo ossessivo-compulsivo[57]. Questo successo è il risultato dell'uso di sistemi avanzati di intelligenza

[55] J. M. STOKES *ET AL.*, "A Deep Learning Approach to Antibiotic Discovery Stokes", *Cell*, Vol. 180, Issue 4, 2020, pp. 688-702, https://pubmed.ncbi.nlm.nih.gov/32084340/; A. TRAFTON, "Using AI, MIT researchers identify a new class of antibiotic candidates", *MIT News*, December 20, 2023, https://news.mit.edu/2023/using-ai-mit-researchers-identify-antibiotic-candidates-1220.

[56] A. MULLARD, "Creating an AI-first drug discovery engine", *Nature Reviews*, 13 September 2024, https://www.nature.com/articles/d41573-024-00149-6

[57] R. J. GEUKES FOPPEN, V. GIOIA, A. ZOCCOLI, "Nuovi farmaci grazie all'AI: ecco le svolte attese nel 2025", *Agenda digitale.EU*, 15 gennaio, 2025, https://www.agendadigitale.eu/sanita/ai-e-genomica-la-nuova-frontiera-della-scoperta-di-farmaci/.

artificiale per la progettazione e ottimizzazione della struttura molecolare del farmaco.

Entro il 2022, almeno centocinquanta farmaci scoperti grazie all'IA erano già in fase di sviluppo, con più di una dozzina di molecole in sperimentazione clinica. Questi successi iniziali[58] delineano una tendenza più ampia: gli strumenti fondati sull'intelligenza artificiale sono oggi impiegati in tutte le fasi della ricerca e sviluppo farmaceutico. Essi contribuiscono non solo all'identificazione dei meccanismi patologici e dei target biologici, ma anche alla progettazione di nuove strutture molecolari e alla previsione della tossicità dei composti.

Grazie all'impiego di modelli di *deep learning*, è possibile suggerire molecole con proprietà farmacologiche ottimali, mentre algoritmi generativi consentono la creazione di strutture chimiche innovative, spesso al di fuori della capacità immaginativa dei chimici umani. Inoltre, l'IA è utilizzata per selezionare e valutare virtualmente i candidati farmaceutici

[58] M. AYERS, M. JAYATUNGA, J. GOLDADER, C. MEIER, "Adopting AI in Drug Discovery", *BGC Blog*, March 29, 2022, https://www.bcg.com/publications/2022/adopting-ai-in-pharmaceutical-discovery.

attraverso simulazioni ad alta fedeltà, riducendo la necessità di sperimentazioni laboratoriali complesse e dispendiose[59].

Complessivamente, sono **capacità che stanno ridefinendo il paradigma operativo dell'industria farmaceutica, abbattendo i costi di sviluppo e riducendo la componente di tentativi ed errori nel processo di scoperta dei farmaci**. Nel lungo periodo, l'accelerazione della ricerca potrebbe garantire ai pazienti l'accesso a nuovi trattamenti con anni di anticipo, trasformando radicalmente la gestione di numerose patologie per le quali attualmente non esistono terapie efficaci.

Oltre agli aspetti tecnico-diagnostici connessi con l'interazione tra IA e paziente, alla distribuzione delle tecnologie di medicina personalizzata, all'automazione dei processi amministrativi e alla gestione ospedaliera, emerge chiaramente un problema legato alla *privacy* e alla sicurezza.

[59] M.VERMA, S. HARDENIA, D. KUMAR JAIN, "Machine learning strategies for drug discovery and development", *IJPSR*, 2025, Vol. 16(5), pp. 1194-1208, https://ijpsr.com/bft-article/machine-learning-strategies-for-drug-discovery-and-development/?view=fulltext.

Infatti, la digitalizzazione dei processi sanitari espone i pazienti a possibili violazioni della *privacy*, le quali non dipendono esclusivamente dalla sicurezza del sistema, ma anche dal quadro normativo di riferimento, riguardante il trattamento dei dati e l'accesso ad essi da parte di soggetti privati e istituzionali. In altri termini, le informazioni relative al paziente - funzionali all'addestramento e al corretto funzionamento dell'IA - sono potenzialmente suscettibili, in assenza di controlli adeguati e di un quadro normativo rigoroso, di essere utilizzati da terze parti. Tutto ciò a prescindere dal suo consenso esplicito o con un consenso poco informato, che nel caso di una mancata anonimizzazione potrebbe condurre a gravi limitazioni concernenti il diritto alla riservatezza e la stessa libertà personale.

Una qualsiasi compagnia assicurativa, a titolo di esempio puramente teorico, se avesse accesso ai dati sanitari di un individuo, potrebbe decidere di negare l'accesso a determinati servizi, sulla base della conoscenza del suo stato di salute potenziale. Per contrastare tali pratiche discriminatorie, è stata

introdotta in Italia la Legge 7 dicembre 2023, n. 193[60,61], che sancisce il diritto all'oblio oncologico. Secondo questa legge, le compagnie assicurative non possono richiedere né utilizzare informazioni relative a patologie oncologiche pregresse se il trattamento attivo si è concluso da più di dieci anni senza recidive, o da più di cinque anni se la diagnosi è avvenuta prima del compimento del ventunesimo anno di età.

Risultano quindi di importanza fondamentale, per la libertà e la sicurezza dei cittadini, oltre che una scrupolosa protezione dei loro dati sanitari e biometrici attraverso i più avanzati strumenti di *cybersecurity*, anche una loro rigorosa anonimizzazione al di fuori della relazione medico-paziente[62]. Tuttavia, per garantire l'avanzamento della ricerca e lo sviluppo di modelli di intelligenza artificiale affidabili, si rende necessario individuare

[60] "Le assicurazioni possono trattare dati sanitari se questi sono necessari per poter fornire servizi previsti nella polizza", 31 maggio 1999, comunicato stampa, GPDP, 31 maggio 1999, https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/48182.

[61] INTESA SANPAOLO PROTEZIONE, Diritto all'Oblio oncologico - Legge 7 dicembre 2023, n.193, 8 maggio 2024, https://www.intesasanpaoloprotezione.com/notizie/oblio-oncologico.

[62] Sulla anonimizzazione dei dati si veda anche il *paper* Aspen Institute Italia, *La gestione del dato*…, op. cit., in particolare il capitolo 2.

soluzioni che consentano l'utilizzo responsabile dei dati personali a fini sperimentali.

In tal senso, si stanno esplorando due strade complementari: da un lato, approcci tecnologici innovativi - come l'impiego di dati sintetici, l'apprendimento federato o tecniche di *privacy* differenziale - che permettano di proteggere l'identità degli individui pur mantenendo l'informatività dei dataset; dall'altro, l'apertura di tavoli regolatori, in collaborazione con l'Autorità Garante per la protezione dei dati personali, finalizzati a definire condizioni normative chiare per l'utilizzo di dati personali nei soli casi in cui sia certificato il reale interesse pubblico della sperimentazione.

I.2.2 FINANZA E SERVIZI BANCARI

Il settore bancario è soggetto a un rigoroso quadro normativo, e la mancata conformità, in particolare nell'ambito della prevenzione del riciclaggio di denaro (Anti Money Laundering - AML), può comportare sanzioni economiche significative. Negli ultimi anni, l'industria bancaria ha subito sanzioni superiori ai

trenta miliardi di dollari[63] per violazioni relative alle normative AML e alle sanzioni finanziarie (cd. *sanctions*). L'intelligenza artificiale sta rivoluzionando le strategie di conformità, consentendo l'automazione del rilevamento delle attività illecite e riducendo il carico di falsi positivi che caratterizza i sistemi tradizionali.

Oltre ai classici approcci supervisionati, in particolare due capacità distintive dell'IA in ambito AML stanno emergendo con forza. La prima è l'uso di tecniche di apprendimento non supervisionato (*unsupervised machine learning*), che permettono di intercettare schemi anomali anche in assenza di etichette o casistiche note, superando la logica reattiva dei soli "*red flag*" tipizzati. La seconda è l'analisi di reti (*network & graph analysis*), che consente di spostare il monitoraggio dal livello della singola transazione a quello delle relazioni tra clienti, gruppi di soggetti o entità giuridiche connesse, rendendo più efficace l'identificazione di attività sospette complesse e distribuite.

I modelli di *machine learning* sono in grado di analizzare milioni di transazioni, identificando schemi sospetti con una precisione

[63] O. HUSAIN , "13 Biggest AML Fines ($500 Million Plus)", 6 May 2024, *Enzuzo*, https://www.enzuzo.com/blog/biggest-aml-fines?utm_source=chatgpt.com

superiore rispetto agli approcci basati su regole fisse. Ad esempio, i sistemi di monitoraggio basati su IA possono esaminare dati eterogenei - cronologia delle transazioni, relazioni tra entità finanziarie, profili dei clienti - per generare punteggi di rischio e individuare in tempo reale clienti o operazioni potenzialmente ad alto rischio. Questo approccio consente a chi si occupa di *compliance* di concentrarsi sulle minacce reali, evitando dispersione di risorse su anomalie prive di rilievo[64,65].

Uno dei problemi principali dei programmi AML tradizionali è l'elevata incidenza di falsi positivi, ossia attività legittime erroneamente classificate come sospette. Questo fenomeno sovraccarica i sistemi di controllo e le risorse umane, rallentando le indagini sui reali rischi di riciclaggio.

[64] M. CARDOSO *ET AL.*, "LaundroGraph: Self-Supervised Graph Representation Learning for Anti-Money Laundering", *arXiv*: 2210.14360, 2022, https://arxiv.org/abs/2210.14360; A. RICADELA, "Anti–Money Laundering AI Explained", Oracle Article, August 28, 2024, https://www.oracle.com/financial-services/aml-ai/.

[65] F. ATTARWALA "Google Cloud Launches AI-Powered Anti-Money Laundering Tool for Banks", *Investopedia*, June 21, 2023, https://www.investopedia.com/google-ai-anti-money-laundering-tool-7550923.

L'adozione di soluzioni basate sull'intelligenza artificiale consente una drastica riduzione degli allarmi ingiustificati, grazie alla capacità dei modelli di apprendere il comportamento normale e identificare solo schemi realmente anomali. Tuttavia, sviluppare modelli efficaci in ambito AML presenta sfide tecniche complesse: da un lato, l'assenza di una variabile target ben definita, che rende difficile l'impiego di metodi supervisionati; dall'altro, la rarità statistica delle transazioni illecite, che complica ulteriormente il bilanciamento dei dataset.

Per affrontare questi vincoli, si ricorre dunque sempre più spesso a tecniche di apprendimento non supervisionato, come il *clustering* e *l'anomaly detection*, che consentono di individuare comportamenti sospetti senza l'ausilio di etichette predefinite. Studi recenti confermano l'efficacia di questi approcci ibridi: ad esempio, il framework ASXAML[66] combina metodi supervisionati e l'ottimizzazione dinamica dei parametri per migliorare la

[66] BAKRY, A.N., ALSHARKAWY, A.S., FARAG, M.S. *ET AL.*, "Automatic suppression of false positive alerts in anti-money laundering systems using machine learning", *J Supercomput* 80, 2024, pp. 6264–6284, https://link.springer.com/article/10.1007/s11227-023-05708-z.

rilevazione delle attività sospette riducendo al tempo stesso i falsi positivi, offrendo un equilibrio tra efficacia e precisione.

Esperienze concrete dimostrano il valore di questi sistemi. Ad esempio, la messa in opera di un sistema AML basato su IA da parte di Google - in collaborazione con una grande banca internazionale - ha ridotto il volume di allarmi, secondo quanto riportato e pubblicato da Google stessa in merito, di oltre il 60%[67], migliorando al contempo la capacità di individuare attività realmente sospette con un incremento di due/quattro volte rispetto ai metodi tradizionali[68].

Un'altra istituzione bancaria europea ha ottenuto una riduzione del 60% dei falsi positivi (con l'obiettivo di raggiungere l'80%) e un aumento del 50% dei rilevamenti effettivi di attività illecite[69].

Sono risultati che dimostrano come l'intelligenza artificiale possa migliorare l'efficacia delle indagini, liberando risorse

[67] R. D. MAY, "Fighting money launderers with artificial intelligence at HSBC", Google Cloud Blog, November 30, 2023, https://cloud.google.com/blog/topics/financial-services/how-hsbc-fights-money-launderers-with-artificial-intelligence.

[68] F. ATTARWALA, *art.cit*.

[69] "Danske Bank Utilises AI to Enhance Fraud Detection", *AI.Business*, March 9, 2024, https://ai.business/case-studies/enhancing-fraud-detection-through-ai-a-danske-bank-journey/.

umane per focalizzarsi sui veri rischi di riciclaggio anziché su anomalie irrilevanti.

Inoltre, l'IA sta velocizzando i processi di conformità: analisi complesse dei flussi di denaro o schemi di riciclaggio complessi, che in passato potevano richiedere settimane, possono ora essere completate in pochi giorni, con un miglioramento complessivo dell'efficienza operativa.

L'integrazione dell'intelligenza artificiale nei processi decisionali finanziari solleva importanti questioni etiche e sociali, in particolare per quanto riguarda la parzialità algoritmica (*algorithmic bias*). Gli strumenti di IA, pur essendo progettati per elaborare decisioni in modo oggettivo e basato sui dati, possono inavvertitamente perpetuare o amplificare i pregiudizi preesistenti nei dataset storici. Tale fenomeno - a cui alcuni settori sono particolarmente soggetti, come risorse umane e giustizia - è critico anche nel settore bancario[70], dove l'intelligenza artificiale è sempre più utilizzata per valutare il merito creditizio, gestire il rischio e determinare l'accesso ai servizi finanziari.

[70] R. KUMAR, "Eliminating AI bias from bank decision-making", BAI Blog, May 26, 2022 https://www.bai.org/banking-strategies/eliminating-ai-bias-from-bank-decision-making.

Un esempio emblematico riguarda l'utilizzo dell'IA nei modelli di concessione del credito. Se le decisioni passate del sistema finanziario presentavano, anche in modo sottile, un pregiudizio nei confronti di determinati gruppi socioeconomici, un modello di *machine learning* addestrato su quei dati potrebbe replicare o amplificare tale distorsione, anche qualora variabili sensibili come genere o etnia non siano esplicitamente incluse nell'analisi. Ciò perché i modelli predittivi identificano correlazioni tra molteplici fattori, alcuni dei quali potrebbero fungere da proxy indiretti per attributi discriminatori.

Un caso di grande risonanza è stato quello della Apple Card nel 2019, una carta di credito emessa da Goldman Sachs[71]. L'algoritmo utilizzato per la valutazione del credito è stato accusato di discriminare le donne, concedendo loro limiti di credito significativamente inferiori rispetto agli uomini con profili finanziari simili[72].

[71] BBC, "Apple's 'sexist' credit card investigated by US regulator", 11 November 2019, https://www.bbc.com/news/business-50365609.

[72] W. KNIGHT, "The Apple Card Didn't 'See' Gender - and That's the Problem", *Wired*, November 19, 2019, https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem.

La controversia ha scatenato un forte dibattito pubblico e ha portato a un'indagine regolatoria sul potenziale pregiudizio di genere del modello impiegato. Goldman Sachs ha difeso il proprio sistema, affermando che l'algoritmo non teneva conto direttamente del genere, eppure l'analisi dei risultati ha suggerito la presenza di una discriminazione indiretta, probabilmente derivante dall'utilizzo di variabili proxy.

Tale episodio ha evidenziato una problematica centrale nell'uso dell'IA nei servizi finanziari: **anche algoritmi apparentemente "neutrali" possono produrre risultati distorti, rendendo complesso il monitoraggio della loro equità**[73].

La neutralità algoritmica, d'altra parte, è un obiettivo difficile da raggiungere in quanto - nella pratica - esisterà sempre un essere umano che sceglie i dati di addestramento e ci saranno sempre state scelte umane che l'IA, allenata a riconoscere questi pattern, cercherà di imitare. Inoltre, l'*audit* di modelli avanzati di *machine*

[73] F. MATTASSOGLIO, "La Corte di giustizia europea, algoritmi e *credit scoring*. L'apertura del vaso di Pandora delle società che si "limitano" a elaborare gli scoring", Note, *Diritto Bancario*, 10 gennaio 2025, https://www.dirittobancario.it/art/la-corte-di-giustizia-europea-algoritmi-e-credit-scoring-lapertura-del-vaso-di-pandora-delle-societa-che-si-limitano-a-elaborare-gli-scoring/.

learning risulta spesso difficoltoso a causa della loro opacità e complessità, sollevando interrogativi sulla responsabilità, la trasparenza e la necessità di regolamentazioni adeguate. In questo contesto, la cosiddetta *explainability* - ossia la capacità di comprendere, tracciare e giustificare le decisioni assunte dall'algoritmo - emerge come una priorità tecnica e regolatoria per garantire la fiducia degli utenti e la conformità ai principi di equità e *accountability*. Tuttavia, raggiungere un adeguato livello di "spiegabilità" è particolarmente complesso per i modelli ad alte prestazioni, come le reti neurali profonde, che operano su strutture altamente non lineari e multilivello, difficili da interpretare anche per gli sviluppatori stessi.

L'intelligenza artificiale sta intervenendo non solo nella razionalizzazione dei processi operativi e nella gestione del rischio, ma anche nella ridefinizione della progettazione, emissione e distribuzione di strumenti finanziari complessi. Tra le applicazioni più avanzate emerge l'impiego della stessa nella creazione di strumenti finanziari personalizzati e conformi ai vincoli normativi delle diverse giurisdizioni. Questo approccio,

che si può definire *AI-Tailored Compliant Financial Instruments*[74], rappresenta un'innovazione di grande rilievo nell'ambito della finanza strutturata, consentendo l'integrazione di forme di finanziamento privato in strumenti custodiali, completamente regolamentati e adattati alle esigenze di investitori professionali e *retail* nei vari mercati globali[75].

L'eterogeneità delle normative finanziarie a livello internazionale rappresenta da sempre un ostacolo significativo alla standardizzazione e scalabilità degli strumenti di investimento, rendendo complessi i processi di emissione, distribuzione e gestione di titoli conformi nei mercati transnazionali.

Tradizionalmente, l'adattamento a regolamentazioni divergenti ha richiesto tempi lunghi e costi elevati, coinvolgendo esperti legali, *compliance officer* e istituzioni finanziarie per garantire il rispetto delle disposizioni locali. L'introduzione dell'IA in questo ambito sta riducendo sensibilmente tali criticità: i **modelli di *machine learning* sono ora in grado di analizzare**

[74] R. AGARWAL, A. KREMER *ET AL.*, "How generative AI can help banks manage risk and compliance", Mc Kinsey, Article,, March 1, 2024, https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/how-generative-ai-can-help-banks-manage-risk-and-compliance.

[75] Secondo la definizione della compagnia Baund, cfr. il loro sito: https://www.baund.ai/.

automaticamente il quadro normativo vigente in diverse giurisdizioni, generando strutture finanziarie conformi e ottimizzate per ciascuna tipologia di investitore.

Un problema centrale è la natura opaca dei sistemi basati su IA, spesso definiti *black box models*[76], poiché i loro meccanismi decisionali possono risultare di difficile interpretazione anche per gli stessi sviluppatori. La complessità di questi algoritmi potrebbe ostacolare la tracciabilità delle decisioni finanziarie, generando preoccupazioni in termini di fiducia, responsabilità e *accountability* regolatoria. Inoltre, **l'adozione su larga scala dell'IA nei mercati finanziari potrebbe amplificare il rischio sistemico**: se più istituzioni utilizzassero modelli predittivi simili, si potrebbero creare comportamenti convergenti che esacerberebbero le fluttuazioni del mercato, riducendo la diversificazione decisionale e aumentando la vulnerabilità collettiva in situazioni di stress.

[76] HASSIJA, V., CHAMOLA, V., MAHAPATRA, A. *ET AL.*, "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence", *Cogn. Comput.* 16, 2024, pp. 45–74, https://link.springer.com/article/10.1007/s12559-023-10179-8.

Per affrontare queste criticità, si stanno sviluppando tecniche di *explainable* AI (XAI), che mirano a rendere più comprensibili le decisioni dei modelli complessi. Alcuni approcci si concentrano sulla generazione di spiegazioni locali, come nel caso di LIME (Local Interpretable Model-agnostic Explanations) o SHAP (SHapley Additive exPlanations), che analizzano l'influenza di ciascuna variabile sull'*output* del modello per un determinato caso.

Altri strumenti cercano invece di costruire modelli surrogati più semplici - come alberi decisionali o reti neurali a bassa profondità - che approssimino il comportamento del modello originale in modo più leggibile. In ambito finanziario, queste soluzioni sono particolarmente utili per migliorare la trasparenza verso clienti e regolatori, e per garantire che i sistemi non discriminino o amplifichino bias preesistenti.

Rimane tuttavia inteso che spiegare in modo completo e affidabile il funzionamento di reti neurali profonde - soprattutto in contesti ad alta dimensionalità e non lineari - è ancora un tema aperto e altamente complesso. L'interazione tra migliaia o milioni di parametri appresi in modo distribuito rende difficile,

se non impossibile, fornire una spiegazione univoca delle scelte del sistema. La ricerca sulla spiegabilità non è solo una questione tecnica, ma anche epistemologica: riguarda il tipo di conoscenza che consideriamo accettabile nei sistemi automatizzati di sostegno alle decisioni.

In particolare, tre rischi principali emergono dall'adozione diffusa dell'IA nella finanza[77,78,79]:

- Sottostima dei rischi di mercato: se i modelli non considerano correttamente gli eventi estremi (*black swans*), potrebbero condurre a valutazioni errate degli asset.
- Effetti amplificatori delle crisi finanziarie: in caso di eventi avversi, il comportamento simultaneo degli algoritmi potrebbe innescare reazioni a catena, accelerando la volatilità.

[77] A.E. MARTIN, "Misurare l'impatto dell'intelligenza artificiale sulla stabilità finanziario", *Network360*, 1 luglio 2024, https://www.ai4business.it/intelligenza-artificiale/misurare-limpatto-dellintelligenza-artificiale-sulla-stabilita-finanziaria.

[78] L. JURICIC, "La Banca d'Inghilterra monitorerà l'uso dell'IA nella finanza per potenziali rischi", articolo sul sito *Investing.com*, 9 aprile 2025, https://it.investing.com/news/economy-news/la-banca-dinghilterra-monitorera-luso-dellia-nella-finanza-per-potenziali-rischi-93CH-2773283.

[79] A. PEDONE, "Economia e Incertezza", articolo sul sito di Tekta, 18 settembre 2024, https://www.tekta.it/economia-e-incertezza/.

- Dipendenza eccessiva dai modelli di IA: un'errata interpretazione dei segnali di mercato da parte degli algoritmi potrebbe generare distorsioni nelle decisioni di investimento.

Per affrontare queste sfide, è essenziale che le istituzioni finanziarie adottino[80,81,82]:

- Standard rigorosi di *AI ethics*, per garantire che i modelli siano equi, privi di distorsioni sistematiche dovuti - ad esempio - a bias che inseriscono elementi di discriminazione nei criteri di scelta e gestiti con opportuni modelli di *governance* che garantiscano la *human agency (human in the loop)*.
- *Framework* di trasparenza, affinché i clienti possano comprendere i criteri decisionali degli algoritmi.

[80] L. MANITTO, "IA per gestione del rischio e strategie di emergenza: guida pratica per PMI", *Network360,* 21 novembre 2024, https://www.agendadigitale.eu/industry-4-0/ai-per-gestione-del-rischio-e-strategie-di-emergenza-guida-pratica-per-pmi/.

[81] "IA nella cybersecurity: Definito e spiegato", scheda Informativa dal sito di FORTINET: https://www.fortinet.com/it/resources/cyberglossary/artificial-intelligence-in-cybersecurity.

[82] "XAI o eXplainable AI - L'Intelligenza Artificiale Spiegabile cos'è e come funziona", scheda informativa sulla piattaforma "Intelligenza Artificiale Italia", https://www.intelligenzaartificialeitalia.net/post/xai-o-explainable-ai-l-intelligenza-artificiale-spiegabile-cos-%C3%A8-e-come-funziona.

- Misure avanzate di *cybersecurity*, per proteggere le informazioni finanziarie da potenziali vulnerabilità.
- Sistemi di gestione della crisi, necessari soprattutto nel caso di emergenze sistemiche.

I.2.3 *CULTURAL HERITAGE*

Nel campo della conservazione e fruizione dei beni culturali l'IA può giocare senza dubbio un ruolo assai rilevante, attraverso la digitalizzazione del patrimonio, di quelle che sono oggi anche definite testimonianze di civiltà. Esse comprendono: opere d'arte, manoscritti, siti archeologici, ma anche l'urbanistica, la ricostruzione virtuale degli spazi cittadini, degli usi e dei costumi del passato, nonché la restituzione dell'originale, quando tempo e usura abbiano deteriorato il bene al punto da renderlo non pienamente fruibile nella sua reale identità costitutiva.

Grazie, dunque, alla combinazione tra realtà virtuale e IA, è ipotizzabile nel prossimo futuro una fruizione immersiva completa sul piano intellettuale, immaginativo ed esperienziale, capace di rinnovare l'approccio ai beni culturali sia in relazione

al grande pubblico, sia in funzione delle esigenze degli studiosi e degli accademici.

Un aspetto particolarmente significativo riguarda il bilanciamento tra realtà virtuale (VR) e realtà aumentata (AR), due tecnologie spesso considerate complementari, ma che rispondono a logiche e finalità differenti.

La VR permette di costruire esperienze completamente immersive, rendendo possibile, ad esempio, l'esplorazione di ambienti perduti o la ricostruzione di eventi storici non più visibili. Tuttavia, proprio per la sua natura pienamente virtuale, questa tecnologia tende a separare l'utente dal contesto fisico reale, proiettandolo in uno spazio altro, simulato.

Al contrario, la realtà aumentata mantiene un ancoraggio diretto con lo spazio autentico, valorizzando l'esperienza *in situ*. Essa consente infatti di sovrapporre livelli digitali informativi, narrativi o visuali all'ambiente circostante, senza alterarne la fruizione diretta. Questa caratteristica è particolarmente rilevante in ambito culturale e turistico, perché preserva la

relazione con l'autenticità del luogo, evitando che la mediazione tecnologica diventi una sostituzione dell'esperienza fisica.

In questa prospettiva, l'adozione dell'AR risulta particolarmente adatta a sostenere un'idea di visita aumentata e consapevole, in cui lo spettatore non abbandona il contesto, ma vi si immerge con strumenti capaci di espandere la comprensione, stimolare la memoria e attivare nuovi livelli di lettura. Tale approccio, sostenuto dall'intelligenza artificiale, può rafforzare il legame tra il patrimonio culturale e i suoi differenti pubblici, senza sottrarre nulla all'emozione del contatto diretto con i luoghi e le testimonianze materiali della storia.

Ovviamente, l'integrazione dell'intelligenza artificiale nelle pratiche di ricostruzione e valorizzazione del patrimonio culturale solleva una serie di questioni critiche, prima fra tutte quella dell'autenticità. L'IA, infatti, è oggi in grado di generare riproduzioni estremamente realistiche di oggetti, ambienti e narrazioni storiche, spesso senza che l'utente finale sia in grado di distinguere tra contenuto originale e contenuto generato. **Ciò**

pone interrogativi rilevanti sulla fedeltà delle ricostruzioni, sulla tracciabilità delle fonti e sull'autorevolezza del contenuto proposto.

Se, da un lato, queste tecnologie offrono l'opportunità di restituire esperienze di visita immersive e accessibili anche a distanza; dall'altro, si corre il rischio che le ricostruzioni - guidate da modelli predittivi e ottimizzazioni statistiche - semplifichino eccessivamente il dato storico, compromettendo la complessità interpretativa e l'integrità culturale dell'opera o del sito. **Senza adeguate forme di verifica e contestualizzazione, l'utilizzo massivo dell'IA potrebbe portare a rappresentazioni fuorvianti o, peggio, a una sorta di "simulacro algoritmico" del passato**: suggestivo, interattivo, però scollegato dalla sua matrice documentaria e critica. In tal senso, si profila il pericolo di trasformare la cultura in intrattenimento immersivo, con logiche più simili al parco tematico che al museo.

Alcune ulteriori problematiche si aggiungono: legate alla creazione *ex novo* o all'alterazione di contenuti culturali tramite IA generativa, in particolare in relazione al diritto d'autore, all'appropriazione indebita di simboli e opere appartenenti a

contesti culturali minoritari o sacri e alla diffusione di informazioni potenzialmente errate, semplificate o non verificabili. L'assenza di un controllo editoriale trasparente e di una chiara indicazione di ciò che è "storicamente fondato" rispetto a ciò che è "ricostruito" rischia di erodere progressivamente il patto fiduciario tra istituzioni culturali e cittadinanza.

Per evitare che le potenzialità tecniche dell'IA producano effetti distorsivi anziché emancipatori, dunque, è necessario sviluppare modelli di *governance* della riproduzione digitale e dell'autorialità algoritmica, coinvolgendo storici, conservatori, informatici e giuristi. In tale prospettiva, il ruolo dell'IA non dovrebbe essere quello di sostituirsi alla ricerca e alla mediazione umana, ma di aiutarne il rigore, amplificarne l'accessibilità e favorire nuove forme di conoscenza partecipata e tracciabile.

Appare di conseguenza fondamentale garantire l'autenticità e la correttezza scientifica di quanto veicolato attraverso l'uso dell'IA, come anche certificare in modo incontrovertibile la corrispondenza tra copia e originale, per cui si renderà necessaria una rigorosa tracciabilità dei dati in associazione

all'uso di *blockchain* e *watermarking* allo scopo di evitare manipolazioni inverosimili e puramente ideologiche del patrimonio storico-culturale.

Una particolare applicazione dell'IA probabilmente si avrà con l'integrazione con la realtà aumentata. Da tempo alcuni siti turistici, monumentali e museali, tra cui per esempio il Franklin Institute di Philadelphia, si sono avviati in questa direzione[83].

L'intelligenza artificiale può potenziare significativamente la realtà aumentata, personalizzando i contenuti in base al profilo del visitatore, riconoscendo in tempo reale gli oggetti o i contesti inquadrati e generando spiegazioni interattive adattate all'interesse dell'utente. Questa integrazione consente di passare da un'esperienza di visita statica a una narrazione dinamica e aumentata, dove l'ambiente fisico si arricchisce di elementi informativi intelligenti. Già nel 2017/2018, infatti, è stata qui sperimentata con un certo successo la AR, in occasione della mostra relativa all'esercito di terracotta, di cui alcuni esemplari sono stati esposti e resi fruibili al grande pubblico.

[83] THE FRANKLIN INSTITUTE, Augmented Reality, https://fi.edu/en/augmented-reality.

Il quadro complessivo appare sostanzialmente in rapida evoluzione, come lo è d'altronde la stessa IA. Si può, tuttavia, presupporre uno sviluppo che investirà la fruizione delle testimonianze di civiltà tanto da remoto attraverso tour virtuali guidati, quanto in presenza. Oltre alle sfide di ordine puramente tecnico che riguardano l'*hardware* utilizzato - che dovrà necessariamente essere aggiornato costantemente in relazione ai rapidi progressi tecnologici in atto - sarà di fondamentale importanza l'aspetto contenutistico.

La vera sfida consisterà nel cercare di evitare l'alterazione sistematica dei dati storici in funzione commerciale e spettacolare, così come la sparizione di contenuti culturali autentici sovrastati dalle performance strumentali a disposizione. Tutte queste innovazioni necessitano, infatti, di essere guidate da un approccio umanistico rigoroso, attraverso il coinvolgimento interdisciplinare di esperti e studiosi dei settori di volta in volta considerati.

Gli specialisti, tra l'altro, dove possibile - onde evitare il più possibile la diffusione sistematica di pregiudizi, bias cognitivi e

stereotipi - andrebbero coinvolti in modo da avere una rappresentanza eterogenea di orientamenti, scuole e metodologie che nei settori umanistici hanno un carattere di oggettività e universalità minore rispetto alle scienze esatte.

Si profila a tal proposito anche il problema dell'uso commerciale dell'IA in relazione alla conservazione e fruizione del patrimonio culturale. In tal senso emergono questioni connesse alla proprietà intellettuale delle opere digitalizzate e all'eventuale compenso da riconoscere a chi ne detiene la proprietà fisica.

Su questo punto non sarà affatto semplice intervenire, tanto per via dell'ipotizzabile frammentazione della digitalizzazione, ovvero l'inevitabile numerosità dei potenziali autori materiali o simbolici delle copie, che a loro volta potranno avvalersi di copie di copie difficilmente controllabili, per non parlare della difficoltà di uniformare i dispositivi legislativi a livello internazionale. Anche la manipolazione - finalizzata a orientare il significato storico e culturale delle opere - sarà difficile da arginare sul piano legislativo, poiché va a sovrapporsi

inevitabilmente alla libertà di pensiero e di espressione, costituzionalmente garantita in tutti i sistemi politici che adottano la formula liberaldemocratica.

Diverso, invece, ma non meno rilevante, il caso dell'automazione e della completa virtualizzazione delle opere che potrebbe avere un impatto negativo sulla guida umana e sull'esperienza diretta. Tuttavia, è importante riconoscere che l'intelligenza artificiale può anche promuovere la democratizzazione della cultura, rendendo l'accesso al patrimonio culturale più inclusivo per persone con disabilità, anziani o coloro che, per motivi economici o logistici, non possono visitare fisicamente i luoghi culturali. Ad esempio, l'uso di assistenti virtuali, traduzioni in tempo reale e sistemi di raccomandazione personalizzati sta trasformando il settore turistico, rendendolo più accessibile e inclusivo per tutti[84].

[84] "Intelligenza artificiale, turismo e accessibilità: una rivoluzione inclusiva", 19 giugno 2024, nota informativa di NODE - Digital Innovation Hub, https://www.dih.node.coop/News/intelligenza-artificiale-turismo-e-accessibilit224-una-rivoluzione-inclusiva.

In tal senso, è stata rilevata una tendenza a breve verso lo sviluppo di un turismo superficiale, centrato in prevalenza su esperienze emozionali prive di autentico spessore culturale, dove le opere sono apprezzate esclusivamente per la loro capacità di stupire e meravigliare, piuttosto che per il loro intrinseco valore storico, artistico e identitario. Un fenomeno analizzato come parte del "post-turismo", in cui il turista è consapevole dell'inautenticità delle attrazioni e le consuma come prodotti di intrattenimento[85].

Porre rimedio a tutte queste tendenze sembra assai complesso, tuttavia, volendo indicare per sommi capi delle linee di azione, esse potrebbero passare per la certificazione delle piattaforme di IA da parte di un'autorità accademica indipendente che ne garantisca la correttezza contenutistica, per l'adozione di licenze *open access* che ne eviterebbero un'eccessiva mercificazione e per modelli di *revenue sharing* tra privati e istituzioni culturali. In ogni caso, la condizione necessaria, anche se non sufficiente, per

[85] M. SMITH *ET AL.*, "Post-Tourism", in *Key Concepts in Tourist Studies*, Sage Publications 2010, https://sk.sagepub.com/dict/mono/key-concepts-in-tourist-studies/chpt/posttourism; NING, W. (2017) "Rethinking authenticity in tourism experience", in *The political nature of cultural heritage and tourism* Routledge, 2017, pp. 469-490, https://doi.org/10.1016/S0160-7383(98)00103-0.

evitare scivolamenti di carattere culturale in senso puramente mercificante, rimane comunque legata al ruolo svolto dalle istituzioni educative come la scuola e la famiglia. Qui si manifesta la necessità e l'urgenza di una pedagogia delle testimonianze di civiltà che, come per altri aspetti dell'esistenza umana nella sua interazione con l'IA, sarà comunque determinante.

I.2.4 SPORT E ANALISI DELLE PERFORMANCE ATLETICHE

L'intelligenza artificiale sta trasformando in modo radicale la preparazione atletica e le strategie di allenamento, permettendo agli atleti di ottimizzare le proprie prestazioni attraverso un approccio fondato sui dati. L'utilizzo di *wearables* intelligenti, come pettorine GPS, sensori biometrici e dispositivi di monitoraggio della performance, consente di raccogliere informazioni in tempo reale su parametri fondamentali quali velocità, frequenza cardiaca, carico di lavoro e dinamica del movimento. Gli algoritmi di IA elaborano questa mole di dati, individuando schemi nascosti e correlazioni che sfuggirebbero all'occhio umano[86].

[86] Cfr. https://www.datacamp.com/blog/ai-in-sports-use-cases.

Ad esempio, un sistema di IA può rilevare una leggera diminuzione della cadenza di un velocista negli ultimi metri di una sessione di allenamento, suggerendo esercizi mirati per migliorare la resistenza. Analogamente, attraverso l'analisi di metriche personalizzate come la variabilità della frequenza cardiaca, i tempi di recupero e la biomeccanica dei movimenti, i modelli di apprendimento automatico possono creare programmi di allenamento personalizzati, adattandoli alle esigenze fisiche e alle caratteristiche tecniche di ogni atleta[87]. È un approccio che consente di superare le fasi di stallo nelle prestazioni con regimi di allenamento ottimizzati, massimizzando l'efficacia del lavoro svolto.

Un ulteriore ambito in cui l'intelligenza artificiale va affermandosi come tecnologia rivoluzionaria è l'analisi della tecnica sportiva. Grazie all'integrazione di *computer vision* e algoritmi di apprendimento automatico, gli allenatori possono ottenere *feedback* immediati sulla biomeccanica degli atleti[88]. Ad

[87] Cfr. https://www.forbes.com/sites/kathleenwalch/2024/08/16/how-ai-is-revolutionizing-professional-sports/.

[88] Cfr. https://www.technogym.com/my/wellness/biomechanics-understanding-the-terms-that-make-our-bodies-move/.

esempio, nel golf, l'IA[89] può analizzare e perfezionare lo *swing* di un giocatore, mentre nel sollevamento pesi può valutare la postura durante l'esecuzione degli esercizi.

Uno studio ha evidenziato poi come l'applicazione di sensori alle macchine per il sollevamento pesi possa fornire dati a sistemi di IA capaci di valutare automaticamente la tecnica dell'atleta, correggendo in tempo reale eventuali inefficienze biomeccaniche[90]. È un tipo di allenatore virtuale che rappresenta un'innovazione significativa, permettendo agli atleti di affinare i propri movimenti con una precisione difficilmente raggiungibile con le metodologie tradizionali.

Nel calcio[91], l'IA è utilizzata per analizzare diverse metriche, un esempio calzante è il calcolo della *minaccia attesa* (*Expected Threat* o xT), ossia i valori che valutano la probabilità che una determinata azione in campo conduca a un gol. Tale approccio supera le tradizionali metriche degli *expected goals* (xG), poiché considera non solo le conclusioni a rete, ma anche le azioni

[89] Cfr. https://www.mad.tf.fau.de/2023/10/19/id-2362/.

[90] Cfr. https://pmc.ncbi.nlm.nih.gov/articles/PMC3761781/.

[91] Cfr. https://www.footballytics.ch/post/expected-threat-xt-the-best-offensive-metric-so-far.

precedenti che aumentano la pericolosità offensiva, come passaggi e *dribbling*. Ad esempio, un passaggio che sposta il pallone in una zona del campo con un'alta probabilità di creare un'occasione da gol riceverà un punteggio xT elevato, riconoscendo così il contributo del giocatore nell'incrementare la minaccia offensiva della squadra[92].

L'adozione di queste tecnologie avanzate consente agli allenatori di identificare e correggere schemi di movimento potenzialmente dannosi. Ad esempio, nel basket[93], l'analisi biomeccanica basata su IA può individuare schemi di atterraggio errati, che esercitano una pressione eccessiva sulle ginocchia, aumentando il rischio di infortuni. Intervenendo preventivamente sulla tecnica di salto, è possibile migliorare le prestazioni e ridurre significativamente il rischio di lesioni, offrendo un doppio vantaggio per atleti e squadre.

[92] Cfr. https://blog.kama.sport/what-is-the-expected-threat-metric-in-football-and-why-it-matters.

[93] Cfr. https://www.folio3.ai/blog/ai-for-biomechanical-analysis/.

Le squadre professionistiche ormai integrano tracker indossabili (GPS, accelerometri, cardiofrequenzimetri), test di *fitness* e persino registri del sonno all'interno di modelli di *machine learning*, che valutano il rischio di infortunio su base giornaliera.

Un esempio emblematico proviene dal calcio professionistico: il *club* Real Salt Lake della Major League Soccer ha adottato la piattaforma basata su IA Zone7[94], un sistema avanzato di monitoraggio che, nell'arco di 26 settimane, ha identificato con precisione i segnali precoci di rischio di infortunio. I risultati sono stati straordinari:

- riduzione del 57% degli infortuni rispetto alla stagione precedente;
- previsione con una settimana di anticipo del 69% degli infortuni effettivamente verificatisi.

Questi dati evidenziano il potenziale rivoluzionario dell'IA nella prevenzione degli infortuni. Il sistema ha individuato segnali precoci - lievi aumenti nei marcatori di fatica, variazioni impercettibili nella biomeccanica del movimento - che sarebbero

[94] Cfr. https://www.sportsbusinessjournal.com/Daily/Issues/2020/06/23/Technology/mls-soccer-real-salt-lake-zone7-injury-risk-prevention/.

potuti sfuggire anche agli staff tecnici più esperti. Gli allenatori descrivono il processo come l'identificazione di "*outlier*": atleti che in apparenza non mostravano problemi, ma che l'IA classificava come ad alto rischio sulla base dell'analisi simultanea di molteplici variabili. Sulla base di queste informazioni, il team medico e lo staff tecnico possono modificare i carichi di lavoro, ridurre l'intensità degli allenamenti o adottare strategie di recupero mirate, prevenendo l'insorgenza di infortuni prima ancora che si manifestino.

L'intelligenza artificiale non solo **sta rivoluzionando la *performance* sportiva, ma sta anche ridefinendo il concetto di benessere e accessibilità allo sport** attraverso un approccio integrato che unisce prevenzione, salute mentale e motivazione.

Per quanto riguarda l'accessibilità[95], l'intelligenza artificiale nello sport favorisce l'inclusione sociale rendendo l'attività fisica più accessibile, personalizzata e adattabile. Tecnologie intelligenti permettono di sviluppare percorsi su misura per

[95] REGIONE LOMBARDIA, *Open Innovation*, "Lo sport come ponte per l'inclusione sociale, tra PNRR e nuove tecnologie - I fondi della Missione 5 Componente 2.3", 13 luglio 2023, https://www.openinnovation.regione.lombardia.it/en/news/news/view?id=7369.

anziani, persone con disabilità o in fase di riabilitazione, abbattendo barriere fisiche e cognitive. In questo modo, l'IA contribuisce a democratizzare lo sport, trasformandolo in uno spazio aperto e inclusivo per tutti.

Per quanto concerne il tema del benessere, si profila un'evoluzione del concetto di *wellness* che possa integrare dati biometrici, IA e programmi individuali onde promuovere uno stile di vita sano e prevenire le patologie. È una visione secondo la quale, l'80% della salute umana dipende da fattori epigenetici, ovvero dallo stile di vita, mentre solo il 20% è determinato dalla genetica. L'*Healthness* si propone quindi come una "medicina preventiva", fondata su strategie scientifiche e *data-driven* per migliorare le prestazioni fisiche e proteggere il benessere futuro[96].

In parallelo, la *gamification* emerge come una potente leva per incentivare l'attività fisica e il benessere. Integrando elementi ludici come punti, sfide, premi e classifiche, la *gamification* rende l'allenamento più coinvolgente e motivante. Un approccio particolarmente efficace nel promuovere l'adozione di stili di vita

[96] Per un esempio di tale paradigma di salute, cfr. https://www.quotidiano.net/economia/technogym-corre-verso-il-futuro-db179b5a.

attivi, migliorare l'adesione ai programmi di esercizio e favorire la formazione di abitudini salutari[97].

Studi recenti hanno dimostrato appunto come la *gamification* possa aumentare la partecipazione alle attività fisiche[98]. Ad esempio, l'integrazione di elementi di gioco nei programmi di *fitness* ha portato a un aumento della frequenza di partecipazione e a un miglioramento dell'impegno degli utenti.

L'integrazione di IA, *Healthness* e *gamification* rappresenta dunque una sinergia potente per promuovere la salute, il benessere e la prevenzione, trasformando l'attività fisica in un'esperienza personalizzata, motivante e sostenibile.

I.2.5 MANIFATTURA

L'intelligenza artificiale è una tecnologia abilitante trasversale, capace di trasformare profondamente non solo i settori ad alta intensità di conoscenza, ma anche comparti produttivi più maturi e tradizionali, come la manifattura. L'adozione dell'IA nel settore manifatturiero sta accelerando a livello globale,

[97] Cfr. https://barcainnovationhub.fcbarcelona.com/blog/benefits-limits-gamification-sports/.
[98] Cfr. https://www.smartico.ai/blog-post/gamifying-sports.

spinta dalla ricerca di nuovi vantaggi competitivi in termini di produttività, qualità, sostenibilità e resilienza delle supply chain. Secondo una stima di McKinsey, l'adozione diffusa dell'IA nei processi produttivi potrebbe generare un aumento del valore aggiunto industriale mondiale fino a 3.000 miliardi di dollari entro il 2030, con impatti significativi lungo tutte le fasi della catena del valore: progettazione, pianificazione, produzione, manutenzione, logistica e post-vendita[99].

Il settore manifatturiero, storicamente al centro delle grandi rivoluzioni industriali - dalla meccanizzazione all'automazione - si trova di fronte a un nuovo passaggio epocale. L'introduzione dell'IA non si limita a sostituire o aggiornare le macchine, ma impone una trasformazione profonda dei processi e dei modelli organizzativi, richiedendo una *governance* attenta del transitorio e un accompagnamento mirato delle persone, chiamate a interagire con sistemi digitali sempre più autonomi, adattivi e intelligenti.

[99] MCKINSEY (ed. by), *The economic potential of generative AI: The next productivity frontier*, Report, June 2023, https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier.

Le applicazioni più mature dell'IA tradizionale - in particolare il *machine learning* supervisionato e non supervisionato - sono oggi già integrate in diversi *Manufacturing Execution Systems* (MES). Questi strumenti consentono, tra l'altro, di prevedere in modo più accurato la produzione, riducendo gli scarti; ottimizzare la gestione del magazzino secondo logiche *just-in-time*; monitorare le prestazioni delle macchine in tempo reale e anticipare guasti critici, grazie alla manutenzione predittiva; classificare e gestire gli scarti industriali favorendo l'economia circolare; migliorare anche l'efficienza energetica degli impianti attraverso la gestione intelligente dei consumi. Il World Economic Forum ha evidenziato come l'IA possa contribuire fino al 20% della riduzione necessaria delle emissioni industriali in Europa al 2030, con un ruolo centrale nella decarbonizzazione della manifattura[100].

A fianco delle applicazioni tradizionali, si stanno affacciando scenari d'uso dell'IA generativa, ancora in fase esplorativa ma di grande potenziale. Questi includono la generazione automatica

[100] WORLD ECONOMIC FORUM, "Digital Tech Can Reduce Emissions by up to 20% in High-Emitting Industries", 24 May 2022, https://www.weforum.org/press/2022/05/digital-tech-can-reduce-emissions-by-up-to-20-in-high-emitting-industries/.

di documentazione tecnica, l'assistenza conversazionale agli operatori in fase di assemblaggio e manutenzione, la simulazione di *concept* industriali o il sostegno alla progettazione mediante *design* generativo. Tuttavia, l'introduzione di questi strumenti pone sfide importanti legate all'affidabilità delle risposte, alla tracciabilità delle fonti, all'*explainability* e alla sicurezza dei sistemi, soprattutto in ambienti produttivi ad alta criticità.

Nonostante il fermento innovativo, la situazione italiana appare a due velocità. Il settore manifatturiero rappresenta circa il 16% del PIL nazionale e oltre il 75% dell'export[101], ma l'adozione sistemica dell'IA resta limitata. Secondo il rapporto *Imprese e ICT* di ISTAT (2024)[102], solo l'8,2% delle imprese italiane ha dichiarato di utilizzare almeno una tecnologia di intelligenza artificiale, contro una media UE del 13,5%. Il divario è ancora più ampio se si guarda all'integrazione di IA con tecnologie come IoT e *cloud*, oggi adottate principalmente da grandi imprese o settori ad alta intensità tecnologica.

[101] ISTAT, *Conti economici nazionali - Anno 2023*, 23 settembre 2024, https://www.istat.it/comunicato-stampa/conti-economici-nazionali-anno-2023/.

[102] ISTAT, *Imprese e Ict - Anno 2024*, 17 gennaio 2025, https://www.istat.it/comunicato-stampa/imprese-e-ict-anno-2024/.

Le principali barriere all'adozione dell'IA nel manifatturiero italiano riguardano: la struttura frammentata del tessuto industriale, con una forte prevalenza di PMI (che hanno una capacità di spesa inferiore e - a volte - non compatibile con gli investimenti necessari); la scarsità di competenze digitali, soprattutto a livello tecnico e intermedio; la difficoltà di accesso a infrastrutture digitali adeguate e la mancanza di strategie industriali di lungo periodo capaci di integrare l'IA nei processi *core*. Un'indagine di Deloitte ha evidenziato che solo il 51,6% dei produttori globali ha oggi una strategia chiara per l'adozione dell'IA[103].

Per superare questi ostacoli e scalare l'impatto dell'IA nel settore manifatturiero, appare necessario un intervento coordinato su più livelli. Sul piano industriale, è cruciale rafforzare la collaborazione tra PMI e grandi imprese in logiche di filiera, facendo leva su piattaforme condivise e dataset interoperabili. A livello istituzionale, come previsto dall'AI Act, si auspica di riorientare gli incentivi pubblici verso l'integrazione reale delle tecnologie nei processi, evitando la dispersione su microprogetti

[103] J. COYKENDALL, K. HARDIN, J. MOREHOUSE, "2025 Manufacturing Industry Outlook", 20 November 2024, Deloitte, disponibile qui: link.

dimostrativi. Inoltre, occorre investire sulla formazione tecnica (ITS, lauree professionalizzanti, apprendistato duale), potenziare i competence center e **promuovere la definizione di standard europei aperti per l'interoperabilità tra sistemi industriali fondati su IA**.

Secondo Accenture, le imprese manifatturiere che hanno già integrato l'IA nelle *operations* riportano un aumento della produttività fino al 20% e una riduzione dei costi fino al 10%, con benefici significativi in termini di resilienza e *time-to-market*[104]. È quindi evidente che l'adozione consapevole dell'IA rappresenta una leva fondamentale per sostenere la competitività del settore manifatturiero italiano, ma la sua efficacia dipenderà dalla capacità di tradurre il potenziale tecnico in cambiamento organizzativo, valorizzando il capitale umano e costruendo un ecosistema industriale inclusivo e interconnesso.

[104] ACCENTURE (a cura di), *Rethinking the course to manufacturing's future*, 2025, https://www.accenture.com/content/dam/accenture/final/accenture-com/document-3/Accenture-Rethinking-The-Course-To-Manufacturings-Future.pdf.

In definitiva, l'intelligenza artificiale può contribuire in modo decisivo a un rilancio sostenibile e innovativo della manifattura italiana, ma **la transizione digitale richiede visione strategica, investimenti continui e un forte coordinamento pubblico-privato**, al fine di evitare che vi sia un tessuto produttivo marginalizzato, incapace di partecipare ai benefici della rivoluzione algoritmica in atto[105].

I.2.6 AGRICOLTURA

L'intelligenza artificiale rivoluziona anche l'agricoltura, trasformandola in un settore sempre più tecnologico e sostenibile. In Italia, l'agricoltura 4.0 ha raggiunto un valore di 2,3 miliardi di euro nel 2024, evidenziando una crescita significativa rispetto ai cento milioni di euro del 2017[106].

Le applicazioni dell'IA in agricoltura sono molteplici e spaziano dalla gestione delle colture alla previsione delle rese. Ad esempio, l'azienda John Deere ha sviluppato macchinari

[105] Per alcuni spunti al riguardo cfr. L. ZANOTTI, "Le applicazioni AI che migliorano la sostenibilità dell'azienda", NETWORK 360, 4 febbraio 2025, https://www.esg360.it/digital-for-esg/le-applicazioni-ai-che-migliorano-la-sostenibilita-dellazienda/.

[106] "Agricoltura 4.0: un valore di 2,3 miliardi di euro", *Fondazione Bruno Kessler Magazine*, 11 febbraio 2025, https://magazine.fbk.eu/it/news/agricoltura-4-0-un-valore-di-23-miliaerdi-di-euro.

avanzati che, grazie all'IA e a tecnologie satellitari, migliorano la produttività e la sostenibilità delle operazioni agricole. Inoltre, l'uso di sensori intelligenti e droni consente un monitoraggio preciso delle condizioni delle colture, ottimizzando l'uso di risorse come acqua e fertilizzanti[107] .

L'IA offre anche opportunità significative per affrontare le sfide legate al cambiamento climatico. Secondo un rapporto del *Financial Times*, l'IA può migliorare le tecniche agricole, aumentando la produttività anche in aree colpite da eventi climatici estremi.

Inoltre, l'uso di modelli predittivi e gemelli digitali consente una gestione più efficiente delle risorse e una maggiore resilienza delle colture.[108]

Tuttavia, l'adozione dell'IA in agricoltura presenta anche alcuni rischi. La dipendenza da tecnologie avanzate potrebbe escludere

[107] ACCENTURE (a cura di), *Innovate Trends and innovations that matter*, February 2024, https://www.accenture.com/content/dam/accenture/final/accenture-com/document-2/Innovation-Booklet-Feb-2024.pdf.

[108] "From bytes to bushels: How gen AI can shape the future of agriculture", dal sito McKinsey, June 10, 2024, https://www.mckinsey.com/industries/agriculture/our-insights/from-bytes-to-bushels-how-gen-ai-can-shape-the-future-of-agriculture; v. anche L. COLBACK, "How we can use AI to create a better society", *Financial Times*, January 23, 2025, https://www.ft.com/content/33ed8ad0-f8ad-42ed-983a-54d5b9eb2d27

i piccoli agricoltori, aumentando le disuguaglianze nel settore. Inoltre, l'uso intensivo di IA può comportare un elevato consumo energetico, potenzialmente annullando i vantaggi ambientali ottenuti.

In Italia, diverse iniziative stanno promuovendo l'integrazione dell'IA nel settore agroalimentare. Ad esempio, la regione Liguria ha lanciato un progetto innovativo per l'olivicoltura che utilizza sensori intelligenti, IA e droni per ottimizzare la produzione e migliorare la qualità dell'olio. Inoltre, JA Europe e JA Italia hanno avviato programmi per formare giovani imprenditori nell'uso dell'IA nella filiera agroalimentare[109].

In conclusione, l'IA rappresenta una leva fondamentale per trasformare l'agricoltura in un settore più efficiente e sostenibile. Tuttavia, è essenziale affrontare le sfide legate all'accessibilità, alla sostenibilità e all'equità per garantire che i benefici dell'IA siano condivisi da tutti gli attori del settore agricolo.

[109] P. DE ANDREIS, "Liguria Region Launches Innovative Olive Farming Project with AI and Smart Sensors", *Olive Oil Times*, April 10, 2025, "https://www.oliveoiltimes.com/production/liguria-region-launches-innovative-olive-farming-project-with-ai-and-smart-sensors/138285.

I.2.7 TELECOMUNICAZIONI E IT

L'integrazione dell'intelligenza artificiale nei settori delle telecomunicazioni e dell'*information technology* (IT) sta ridefinendo profondamente le dinamiche operative e strategiche di queste industrie. L'adozione di soluzioni basate su IA consente non solo di ottimizzare le performance delle reti e dei sistemi informatici, ma anche di affrontare sfide storiche legate alla gestione di infrastrutture *legacy* e alla produttività degli sviluppatori.

Nel comparto delle telecomunicazioni, l'IA è impiegata per migliorare l'efficienza operativa, la qualità del servizio e l'esperienza del cliente. Ad esempio, un operatore latino-americano ha registrato un incremento del 25% nella produttività degli agenti del *call center* grazie all'adozione di soluzioni di IA generativa, che forniscono raccomandazioni in tempo reale, migliorando le competenze e le conoscenze degli operatori.

Inoltre, l'IA facilita la manutenzione predittiva delle infrastrutture, l'ottimizzazione del traffico di rete e la personalizzazione dei servizi offerti agli utenti[110].

Un ulteriore ambito di applicazione riguarda i *digital twins* delle reti, modelli virtuali che replicano il comportamento delle infrastrutture fisiche. L'integrazione di IA generativa in questi modelli consente di simulare scenari complessi, prevedere anomalie e ottimizzare le operazioni di rete in tempo reale. Ad esempio, Chunghwa Telecom ha utilizzato un *digital twin* potenziato da IA per prevedere la congestione della rete durante eventi ad alta densità di traffico, migliorando la capacità della rete del 14% e mantenendo elevate velocità di *download*[111] .

Nel settore IT, l'IA sta rivoluzionando la produttività degli sviluppatori e la modernizzazione dei sistemi *legacy*. Strumenti come GitHub Copilot, basati su modelli di linguaggio avanzati, assistono gli sviluppatori nella scrittura del codice, suggerendo

[110] "How generative AI could revitalize profitability for telcos", February 21, 2024, sito di McKinsey: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/how-generative-ai-could-revitalize-profitability-for-telcos.

[111] "Chunghwa Telecom and Ericsson harness AI and digital twins to secure network performances during data surge", Press release, Ericsson, February 20, 2025, https://www.ericsson.com/en/press-releases/2/2025/2/chunghwa-telecom-and-ericsson-harness-ai-and-digital-twins-to-secure-network-performances-during-data-surge.

automaticamente frammenti di codice e riducendo il tempo necessario per lo sviluppo e la manutenzione delle applicazioni. Secondo McKinsey, l'adozione di tali strumenti può aumentare la produttività degli sviluppatori del 20-45%, liberando risorse per attività più strategiche e creative[112].

Un aspetto particolarmente rilevante riguarda la modernizzazione dei codebase sviluppati in linguaggi obsoleti come COBOL o LISP, ancora ampiamente utilizzati in settori come la finanza e la pubblica amministrazione. La conversione di questi sistemi in linguaggi moderni rappresenta una sfida significativa a causa dei costi e dei rischi associati. Tuttavia, l'IA offre strumenti avanzati per l'analisi e la conversione automatica di questi codebase. Ad esempio, IBM ha sviluppato watsonx Code Assistant for Z[113], una soluzione basata su IA generativa che assiste nella conversione di applicazioni COBOL in Java, facilitando la transizione da ambienti *mainframe* a piattaforme più moderne.

[112] "What's the future of generative AI? An early view in 15 charts", August 25, 2023, dal sito di McKinsey: https://www.mckinsey.com/featured-insights/mckinsey-explainers/whats-the-future-of-generative-ai-an-early-view-in-15-charts.

[113] "IBM watsonx accelera su COBOL", redazione LineaADP, 5 settembre 2023, https://www.lineaedp.it/featured/ibm-watsonx-accelera-su-cobol/.

Nonostante i benefici, l'adozione dell'IA presenta anche sfide significative. Le preoccupazioni riguardanti la *privacy* dei dati, la sicurezza e l'etica dell'IA sono particolarmente rilevanti nei settori delle telecomunicazioni e dell'IT, dove la gestione di informazioni sensibili è quotidiana. Inoltre, la carenza di competenze specializzate in IA rappresenta un ostacolo alla messa in opera efficace di queste tecnologie. È quindi essenziale investire nella formazione e nello sviluppo di competenze per garantire una transizione fluida verso un ecosistema digitale potenziato dall'IA.

In conclusione, l'intelligenza artificiale offre opportunità straordinarie per trasformare le telecomunicazioni e l'*information technology*, migliorando l'efficienza, la produttività e la qualità dei servizi. Per sfruttare appieno questi benefici, è fondamentale affrontare le sfide associate all'adozione dell'IA attraverso strategie mirate, investimenti in competenze e una *governance* attenta, assicurando che l'innovazione tecnologica sia al servizio di un progresso sostenibile e inclusivo.

I.2.8 DIFESA E SICUREZZA

Le tecnologie basate sull'intelligenza artificiale vanno acquisendo un ruolo sempre più rilevante nei sistemi di difesa e sicurezza, in risposta alla crescente complessità e multidimensionalità delle minacce contemporanee. Lo sviluppo di strumenti digitali capaci di elaborare rapidamente grandi quantità di dati, anticipare anomalie e affiancare decisioni tattiche e strategiche ha favorito un'accelerazione nell'adozione di soluzioni IA da parte delle forze armate, dei ministeri della difesa e delle agenzie di *intelligence* a livello globale (esemplificativo, in questo senso, il recente Atto d'intesa tra Difesa e Agenzia per la Cybersicurezza Nazionale)[114].

Negli Stati Uniti, il Dipartimento della Difesa ha lanciato il progetto "Thunderforge" in collaborazione con *Scale* AI, con l'obiettivo di integrare modelli linguistici avanzati per coadiuvare la pianificazione di missioni e lo spostamento di assetti operativi. Sviluppato anche grazie al contributo di

[114] Atto d'intesa tra Difesa e Agenzia per la Cybersicurezza Nazionale, Roma 9 giugno 2025, pubblicato sul sito del Ministero della Difesa: https://www.difesa.it/primopiano/atto-dintesa-tra-difesa-e-agenzia-per-la-cybersicurezza-nazionale/72639.html. In particolare, definisce funzioni e compiti della "Struttura di Collegamento della Difesa", che opererà in sinergia con l'Agenzia per rafforzare la capacità nazionale di risposta e prevenzione nel dominio cibernetico.

Microsoft e Google, il sistema è concepito per fornire raccomandazioni strategiche ai comandanti militari, elaborando in tempo reale dati raccolti da sensori, *intelligence* e fonti aperte[115]. Si tratta di una delle prime attuazioni operative dell'IA generativa in ambito difensivo, in grado di affiancare il personale nella lettura e sintesi di informazioni critiche.

Parallelamente, il Progetto Maven - attivo dal 2017 - continua a essere un riferimento per l'integrazione di IA e *machine learning* nell'analisi video e immagini satellitari, con finalità di sorveglianza e identificazione automatica di obiettivi sul campo[116]. In Europa, aziende come Helsing AI[117] stanno sviluppando interfacce che aggregano feed da sensori e li restituiscono come visualizzazioni in tempo reale alle forze terrestri e navali. La capacità di rendere intellegibili ambienti

[115] G. DE VYNCK, "Pentagon signs AI deal to help commanders plan military maneuvers, *The Washington Post*, March 5, 2025 https://www.washingtonpost.com/technology/2025/03/05/pentagon-ai-military-scale/.

[116] Lemma *Project Maven* da Wikipedia https://en.wikipedia.org/wiki/Project_Maven. Sul sito del dipartimento della Difesa statunitense le notizie son ferme al 2017. Dal 2022 il progetto è affidato a *National Geospatial Agency*, https://www.defenseone.com/technology/2022/04/nga-will-take-over-pentagons-flagship-ai-program/366098/.

[117] Si veda il sito della *defence company* Helsing: https://helsing.ai/.

operativi complessi è considerata un *asset* critico per garantire superiorità informativa e tempi di reazione più rapidi[118,119].

In questo quadro, il ricorso all'IA consente anche di affrontare nodi strutturali legati all'obsolescenza tecnologica. Molti sistemi in uso nei ministeri e nei comandi militari si basano su *software* sviluppati in linguaggi oggi marginali, come COBOL, la cui manutenzione è onerosa. L'adozione di IA per la conversione automatica dei *codebase legacy* sta già mostrando risultati concreti. L'U.S. Army, ad esempio, ha avviato una profonda revisione digitale, con l'obiettivo di ridurre costi ed errori, razionalizzare le piattaforme e accelerare i processi amministrativi. Un caso citato dal *New York Post* (aprile 2025) mostra come l'uso dell'IA abbia permesso di aggiornare in una sola settimana le *job descriptions* di oltre 300.000 dipendenti civili, attività che prima richiedeva mesi[120].

[118] S. MUKHERJEE, M. KAHN, E. HOWCROFT, "Mission before money: how Trump and Ukraine are helping Europe's defence industry lure AI talent", *Reuters*, April 30, 2025, https://www.reuters.com/business/mission-before-money-how-europes-defence-startups-are-luring-ai-talent-2025-04-30.

[119] M. MEAKER, "A Battlefield AI Company Says It's One of the Good Guys", *Wired*, July 20, 2023, https://www.wired.com/story/helsing-ai-military-defense-tech/.

[120] R. KING, "Army quietly pursues massive digital overhaul expected to save at least $89M: 'Dealing with stuff that was for 25 years ago'", *New York Post*, April 29, 2025,

La rilevanza di queste trasformazioni si riflette anche nelle dinamiche del mercato. Secondo un'analisi di *GlobeNewswire* (novembre 2024)[121], il mercato globale dell'IA in ambito difesa raggiungerà i 18,6 miliardi di dollari entro il 2029, con un tasso di crescita annuale superiore al 30%. La spinta è trainata da investimenti governativi crescenti, dal ritorno della competizione geopolitica su scala globale e dall'emergere di scenari ibridi in cui le minacce *cyber*, cognitive e cinetiche si sovrappongono.

In ambito *cybersecurity*, l'IA rappresenta una leva fondamentale per l'analisi in tempo reale di anomalie nei flussi di rete, per la risposta automatica a incidenti e per l'identificazione di vulnerabilità nei sistemi critici. La combinazione di modelli predittivi e reti neurali permette oggi di anticipare comportamenti malevoli[122] con maggiore accuratezza, contribuendo alla difesa delle infrastrutture civili e militari in un contesto sempre più

https://nypost.com/2025/04/29/us-news/army-quietly-pursues-digital-overhaul-expected-to-save-89m/.

[121] "Artificial Intelligence in Aerospace & Defense Strategic Research Report 2025: Global Market to Reach $44.1 Billion by 2030", *Globe Newswire*, March 17, 2025, https://www.globenewswire.com/news-release/2025/03/17/3043851/28124/en/Artificial-Intelligence-in-Aerospace-Defense-Strategic-Research-Report-2025-Global-Market-to-Reach-44-1-Billion-by-2030-Growing-Adoption-in-Surveillance-Threat-Detection-Operationa.html.

[122] "What Is the Role of AI in Threat Detection?", scheda informativa a cura di Paloalto Networks, https://www.paloaltonetworks.com/cyberpedia/ai-in-threat-detection.

digitalizzato. Il ruolo dell'IA è altrettanto centrale nella protezione delle supply chain digitali, dove la crescente interconnessione espone a rischi di compromissione a monte, lungo l'intero ciclo di vita di *hardware* e *software*.

L'Europa, pur con una spesa militare inferiore rispetto agli Stati Uniti, ha registrato una crescita significativa dell'attenzione verso tali tecnologie[123]. I fondi del programma EDF (European Defence Fund) stanno progressivamente finanziando lo sviluppo di soluzioni IA in ambito militare *dual-use*, con particolare attenzione alla sovranità tecnologica, all'interoperabilità tra forze armate e alla cooperazione industriale transfrontaliera.

Accanto alle opportunità, l'intelligenza artificiale applicata alla difesa solleva interrogativi etici, giuridici e geopolitici. Il dibattito internazionale sullo sviluppo e l'uso di sistemi d'arma autonomi - dai droni fino alle piattaforme terrestri - è tuttora aperto. Fondamentali preoccupazioni riguardano il principio del

[123] T. BRADSHAW and S. PFEIFER, "VC funding in European defence and security tech surges to record $5.2bn", *Financial Times*, February 12, 2025, https://www.ft.com/content/6c21daac-1a07-4fe2-bd32-7237a8285717.

controllo umano significativo[124] ("*meaningful human control*"), l'attribuzione delle responsabilità in caso di errore, l'aderenza al diritto internazionale umanitario e la possibilità che questi strumenti possano alterare l'equilibrio della deterrenza.

In questo contesto, la NATO ha avviato una riflessione sulla definizione di standard etici condivisi per l'uso dell'IA in ambito militare. Il rapporto 2024 della NATO (NATO PA, Report 058 STC[125]) raccomanda l'adozione di principi chiari su trasparenza, robustezza, affidabilità e *accountability* dei sistemi, affermando che l'innovazione tecnologica debba restare coerente con i valori democratici dei paesi membri.

In sintesi, l'intelligenza artificiale sta rafforzando le capacità difensive e strategiche degli attori statali, abilitando una nuova generazione di strumenti che trasformano il ciclo decisionale, le operazioni sul campo e la sicurezza delle infrastrutture critiche. Tuttavia, la crescente integrazione di sistemi autonomi impone

[124] L. TRABUCCO, "What is Meaningful Human Control, anyway? Cracking the Code on Autonomous Weapons and Human Judgment", 21 September 2023, articolo sul sito del Modern War Institute, https://mwi.westpoint.edu/what-is-meaningful-human-control-anyway-cracking-the-code-on-autonomous-weapons-and-human-judgment/.

[125] *NATO and artificial intelligence: navigating the challenges and opportunities*, special report NATO by S.CLEMENT, 24 November 2024, https://www.nato-pa.int/document/2024-nato-and-ai-report-clement-058-stc.

anche un ripensamento delle dottrine, dei controlli e della cooperazione multilaterale.

Per le democrazie, la sfida consiste nel governare questa transizione in modo attivo, garantendo la superiorità tecnologica nel rispetto dei principi di legittimità, trasparenza e responsabilità.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

II.

AVANZAMENTO TECNICO DELL'IA

Con il contributo di

Alessandro Golkar, Fabio Bresciani, Edoardo Degli Innocenti

II.1. INTRODUZIONE A NUOVE TECNOLOGIE EMERGENTI DI INTELLIGENZA ARTIFICIALE

Nel corso del 2024 sono avvenuti progressi significativi in aree chiave come l'IA generativa, l'IA collaborativa e le tecnologie per la *privacy-preserving AI*. La conferma del successo dei *Large*

Language Models (come il recente DeepSeek-V3[126], modello cinese che ha generato particolare clamore sul mercato dati i costi di *training* di due ordini di grandezza inferiore ai concorrenti americani) ha segnato una svolta nell'adozione dell'IA nella società moderna. **Il continuo miglioramento dell'IA ha innescato un'adozione su larga scala in applicazioni quali il servizio clienti, la creazione di contenuti e l'analisi dei dati**. Le aziende stanno rapidamente integrando l'IA generativa per migliorare l'efficienza operativa, sostenere i dipendenti nelle attività ripetitive e sviluppare nuove strategie di interazione con i clienti.

Nell'adozione delle nuove tecniche di IA, le aziende si sono scontrate con le problematiche di *privacy* dei dati e dell'avvicendamento del lavoro umano con quello delle macchine. Nella fase attuale di sviluppo della tecnologia, l'enfasi è posta sull'utilizzo dell'IA come complemento delle attività lavorative esistenti. L'IA collaborativa lavora in simbiosi con gli esseri umani per migliorarne la produttività e la qualità delle decisioni.

[126] DeepSeek API Docs, Introducing DeepSeek-V3, scheda informative, 26 December 2024, https://api-docs.deepseek.com/news/news1226.

In ultimo, le tecnologie di *privacy-preserving AI* mirano a conciliare l'uso massivo dei dati con la protezione della riservatezza e della sicurezza. L'applicazione di approcci in tale direzione come il *federated learning* (approccio ben noto, ad esempio nel *training* delle macchine autonome così come proposto da Google[127]), il *multiparty computation* e l'utilizzo di dati sintetici consente alle aziende di sfruttare l'IA senza esporre informazioni sensibili, favorendo l'adozione in settori regolamentati come la finanza e la sanità.

II.1.1 Intelligenza Artificiale Generativa

L'IA generativa (GenAI) rappresenta una delle aree più dinamiche dell'intelligenza artificiale, con modelli capaci di creare contenuti multimediali di qualità sempre più elevata. Alla base di queste tecnologie vi sono reti neurali profonde (*deep neural networks*), strutture computazionali ispirate al funzionamento del cervello umano, in grado di apprendere rappresentazioni complesse dei dati attraverso molteplici strati di elaborazione. Un ruolo centrale è svolto dai modelli *transformer*, una particolare architettura introdotta nel 2017,

[127] "Federated Learning", an online comic from Google AI, https://federated.withgoogle.com/.

capace di elaborare sequenze di dati (come testi o immagini) in parallelo, oggi alla base della maggior parte dei modelli generativi. In alcuni casi, le prestazioni vengono ulteriormente affinate grazie all'apprendimento per rinforzo (*reinforcement learning*), una tecnica che permette al modello di migliorare le proprie risposte tramite feedback, premiando i risultati considerati più utili o aderenti al contesto d'uso.

Tali modelli, come il noto Midjourney, ma anche DALL·E, Stable Diffusion e Runway, che stanno rivoluzionando settori quali il *marketing*, il *design* e la produzione di contenuti. La capacità di generare testi, immagini, video e suoni realistici apre nuove opportunità, ad esempio, nel campo pubblicitario e nella personalizzazione dei prodotti.

Nel *marketing*, le aziende stanno sfruttando GenAI per generare testi pubblicitari, post per i social media e contenuti personalizzati per le campagne, riducendo i tempi di produzione e migliorando l'*engagement* del pubblico. Si assiste così alla comparsa della prima generazione di influencer artificiali, figure interamente generate da modelli di IA, impiegate per la promozione di prodotti e servizi in ambito

retail, moda e tecnologia. Nella produzione multimediale, strumenti avanzati come Runway Gen-2 o Synthesia consentono di realizzare spot pubblicitari, trailer cinematografici e materiali formativi con costi e tempi inferiori rispetto al passato.

II.1.2 Intelligenza Artificiale Collaborativa

L'IA collaborativa è costituita dai sistemi che lavorano sinergicamente agli operatori umani o insieme ad altri sistemi di IA per migliorare l'efficienza e la qualità delle decisioni. Questi sistemi si basano su algoritmi di apprendimento federato, modelli multi-agente e interfacce uomo-macchina avanzate. Un tratto distintivo e sempre più centrale di questi approcci è la logica *human-in-the-loop* (HITL), che prevede la presenza attiva, consapevole e costante dell'essere umano nei processi decisionali sostenuti dall'IA.

Tale paradigma non rappresenta soltanto un'opzione tecnologica, ma una condizione fondamentale per assicurare che l'intelligenza artificiale operi in modo trasparente, controllabile e in linea con valori umani, in particolare nei contesti ad alta criticità come la salute, la sicurezza e la finanza.

La logica HITL permette di integrare la capacità computazionale dei modelli con l'intuizione, il giudizio esperto e la responsabilità dell'operatore umano. In questo senso, l'essere umano non è solo un controllore o un correttore a valle, ma parte integrante del ciclo decisionale: supervisiona, interviene, apprende e migliora insieme al sistema. Inoltre, tale approccio incrementa l'accettabilità dell'IA in ambito organizzativo, riducendo la resistenza al cambiamento e rafforzando il senso di fiducia, sia per gli operatori che per gli utenti finali.

Le applicazioni spaziano dalla sanità - come già anticipato nel cap. I (§ I.2.1) - dove l'IA coadiuva i medici nella diagnosi e nella personalizzazione dei trattamenti, alla manifattura avanzata, in cui robot intelligenti collaborano con gli operai per ottimizzare la produzione[128].

Un esempio significativo di messa in uso di algoritmi di IA in ambito diagnostico è l'adozione di algoritmi di IA presso il Lahey Hospital & Medical Center negli Stati Uniti, designato

[128] R. SMITH, "Collaborative AI: The Next Frontier," *AI Journal*, vol. 58, no. 2, pp. 45-60, 2023.

come ACR® Recognized Center for Healthcare-AI (ARCH-AI[129]). Qui, un sistema di IA combinato con un *software* di orchestrazione del flusso di lavoro analizza le immagini e assegna priorità ai casi con potenziali risultati critici, come embolie polmonari, emorragie intracraniche e fratture cervicali. **Il sistema non sostituisce il medico, ma lo affianca nel processo decisionale**, evidenziando i casi a rischio e lasciando la valutazione finale all'esperienza clinica dell'operatore. Il medico, in tal senso, rimane al centro del processo, con l'IA che agisce come acceleratore cognitivo.

Nel settore industriale, gli assistenti IA sono integrati nei processi produttivi per monitorare le prestazioni delle macchine e prevedere guasti prima che si verifichino, riducendo i tempi di inattività e migliorando l'efficienza operativa. Un esempio nel settore aerospaziale è rappresentato da *Airbus Skywise*[130], una piattaforma di analisi dei dati specificamente sviluppata per l'industria aeronautica da Palantir a partire dal 2017 e di recente

[129] ACR Recognized Center for Healthcare-AI (ARCH-AI), The first national artificial intelligence (AI) quality assurance program, cfr. https://www.acr.org/Data-Science-and-Informatics/AI-in-Your-Practice/ARCH-AI.

[130] La piattaforma dati Skywise Core [X] di Airbus Aircraft è disponibile qui: https://aircraft.airbus.com/en/services/enhance/skywise-data-platform/skywise-core-x.

aggiornata per far uso delle moderne tecniche di intelligenza artificiale. Skywise raccoglie e analizza enormi quantità di dati provenienti da sensori a bordo degli aeromobili, dai registri di manutenzione e dalle operazioni di volo. Utilizzando algoritmi di *machine learning*, Skywise identifica schemi e anomalie che potrebbero indicare problemi imminenti, consentendo una manutenzione predittiva degli aeromobili. Anche in questo caso, l'IA coadiuva l'ingegnere nella pianificazione degli interventi, ma la decisione finale resta responsabilità dell'essere umano, che valuta le implicazioni operative e di sicurezza.

Tale capacità di previsione aiuta le compagnie aeree a ridurre i tempi di inattività non pianificati, ottimizzare la gestione delle risorse e migliorare l'efficienza complessiva delle operazioni. Ad esempio, Vietnam Airlines utilizza Skywise Predictive Maintenance per aumentare la sicurezza e la sostenibilità della sua flotta, mentre Qantas e Jetstar Airways lo utilizzano per minimizzare i ritardi e ridurre gli incidenti di Aircraft on Ground (AOG)[131] .

[131] "Qantas and Jetstar Airways to optimise operations with Skywise Predictive Maintenance", 26 February 2024, Press Release Airbus Aircraft, https://aircraft.airbus.com/en/newsroom/press-releases/2024-02-qantas-and-jetstar-airways-to-optimise-operations-with-skywise?t.

Nel campo della logistica, un esempio concreto di ottimizzazione attraverso l'IA è il progetto Warehouse Intelligence messo in opera da LPP (azienda attiva nel settore della moda) in collaborazione con PSI Logistics[132]. LPP ha integrato l'intelligenza artificiale nel proprio sistema di gestione del magazzino PSIwms, ottenendo miglioramenti significativi nella gestione dell'inventario e nell'efficienza operativa.

Grazie all'intelligenza artificiale, è stato possibile ottimizzare i percorsi di prelievo delle merci, analizzando i movimenti dei magazzinieri esperti per identificare le strategie più efficienti. Il sistema apprende dai comportamenti umani, genera suggerimenti e migliora nel tempo attraverso un ciclo continuo di feedback: l'essere umano, in questo contesto, non è solo "nel *loop*", ma "al centro del *loop*". Ciò ha portato a un aumento del 30% nella produttività dei prelievi orari. Inoltre, il sistema ha migliorato la gestione dell'inventario, riducendo i rischi di esaurimento scorte e sovra-stoccaggio, ottimizzando così la disponibilità delle merci.

[132] G. SISINNA, "Come l'AI trasforma il *Warehouse Management System*", *AI.4.Business*, 5 marzo 2024, https://www.ai4business.it/intelligenza-artificiale/come-lai-trasforma-il-warehouse-management-system/?t.

II.1.3 TECNOLOGIE PER LA *PRIVACY-PRESERVING AI*

L'evoluzione dell'IA richiede approcci innovativi per garantire la sicurezza e la *privacy* dei dati. In questo contesto emergono tecnologie come il *Federated Learning* (FL), il *Multiparty Computation* (MPC) e i Dati Sintetici.

Federated Learning consente l'addestramento di modelli di IA su dati decentralizzati senza trasferirli in un unico *repository* centrale. Questo approccio migliora la *privacy* e la conformità a normative come il GDPR, permettendo alle aziende di addestrare modelli di IA senza condividere dati sensibili tra diversi soggetti. Nel settore bancario, lo FL potrebbe essere utilizzato per la rilevazione delle frodi, consentendo alle istituzioni finanziarie di identificare schemi sospetti senza esporre i dati dei clienti. In ambito sanitario, lo stesso permette di sviluppare modelli diagnostici, utilizzando dati provenienti da più ospedali senza doverli centralizzare, migliorando così la qualità delle previsioni cliniche.

Multiparty Computation è una tecnica crittografica che permette a più parti di eseguire calcoli su dati condivisi senza rivelarne il contenuto. Questa tecnologia è particolarmente utile nelle

collaborazioni interaziendali, ad esempio nel settore assicurativo, dove più compagnie possono calcolare i rischi congiuntamente senza dover condividere i dati individuali dei clienti. Analogamente, nelle operazioni finanziarie, lo stesso facilita la condivisione di analisi di mercato tra banche senza compromettere la riservatezza delle informazioni.

Dati sintetici sono informazioni generate artificialmente mediante algoritmi matematici o modelli di intelligenza artificiale, progettati per replicare le caratteristiche statistiche dei dati reali, senza contenere riferimenti diretti a individui o entità esistenti. Questi dati sono utilizzati in vari ambiti, come l'addestramento di modelli di *machine learning*, il *testing* di *software* e la ricerca, offrendo una soluzione efficace per superare le limitazioni legate alla disponibilità di dati reali e alla protezione della *privacy*. Tuttavia, la generazione e l'utilizzo dei dati sintetici presentano sfide significative. È fondamentale trovare un equilibrio fra tre aspetti chiave:

Accuratezza: i dati sintetici devono mantenere le proprietà statistiche dei dati reali per garantire che le analisi e i modelli sviluppati su di essi siano validi e generalizzabili.

Utilità: i dati devono essere sufficientemente dettagliati e rappresentativi per essere utili nelle applicazioni previste, come l'addestramento di modelli predittivi o l'analisi statistica.

Privacy: è essenziale che i dati sintetici non permettano la re-identificazione di individui o la divulgazione di informazioni sensibili, rispettando le normative sulla protezione dei dati.

La difficoltà sta nel bilanciare questi tre elementi: aumentare l'accuratezza e l'utilità può comportare rischi per la *privacy*, mentre rafforzarla può ridurre l'utilità e l'accuratezza dei dati[133].

Nel settore sanitario, i dati sintetici sono utilizzati per addestrare algoritmi di diagnostica su malattie rare senza compromettere la riservatezza dei pazienti[134].

Nel settore finanziario, permettono di testare modelli di rischio senza esporre dati sensibili[135].

[133] LAUTRUP, A.D. ET AL., "Syntheval: a framework for detailed utility and privacy evaluation of tabular synthetic data". Data Min Knowl Disc 39, 6 (2025), https://arxiv.org/abs/2404.15821.

[134] GIUFFRÈ, M., SHUNG, D.L., "Harnessing the power of synthetic data in healthcare: innovation, application, and privacy", *npj Digit. Med*. 6, 186 (2023), https://www.nature.com/articles/s41746-023-00927-3.

[135] *Synthetic Data: Key Use Cases*, Report ed. by Reply, s.d., https://www.reply.com/en/data-world/synthetic-data-business-opportunities-and-use-cases.

Nel *retail*, i dati sintetici aiutano le aziende a sviluppare strategie di *marketing* predittivo basate su simulazioni di comportamento dei consumatori[136].

II.1.4 IA GENERALE (AGI)

L'Intelligenza Artificiale Generale (*Artificial General Intelligence*, in sigla AGI) rappresenta uno degli obiettivi più ambiziosi della ricerca nel campo IA. A differenza degli attuali sistemi, che rientrano nel dominio dell'IA ristretta (*narrow AI*) e sono progettati per svolgere specifici compiti con grande efficienza, l'AGI punta a sviluppare agenti capaci di apprendere, ragionare e adattarsi autonomamente a una vasta gamma di contesti e problemi, con una flessibilità cognitiva simile a quella umana[137].

Come illustrato anche nel noto schema comparativo proposto da Deeptech Center[138], esistono **sei livelli progressivi nell'evoluzione dell'IA**: i sistemi attuali - definiti

[136] M. HENRY, "The Rise of Synthetic Data in Retail Automation: A Technical Exploration", Neurolabs, 5 June 2024, https://www.neurolabs.ai/post/the-rise-of-synthetic-data-in-retail-automation-a-technical-exploration.

[137] D. BERGMANN, C. STRYKER, "What is artificial general intelligence (AGI)?", IBM TOPICS, 17 September 2024, https://www.ibm.com/think/topics/artificial-general-intelligence.

[138] G. PERSCHEID, "Innovative AI - Artificial Intelligence (AI) and its impact on businesses", Innovative AI, March 8 2024, https://news.deeptechcenter.org/p/innovative-ai-artificial-intelligence-451.

"ristretti" - sono molto efficaci, ma operano solo all'interno di confini funzionali ben delimitati; al livello intermedio si colloca l'AGI, che mira a generalizzare la conoscenza e ad affrontare problemi nuovi senza riaddestramento; infine, l'orizzonte più remoto è rappresentato dall'IA superintelligente (in sigla, ASI), una tecnologia ipotetica che supererebbe l'intelligenza umana in ogni dimensione cognitiva. Lo schema concettuale suggerisce come l'AGI rappresenti un punto di transizione critico: un cambio di paradigma, non solo incrementale, tra macchine "specializzate" e sistemi veramente "generali".

Uno degli aspetti centrali della AGI è proprio questa capacità di trasferire conoscenza tra domini differenti, cosa che i modelli attuali non sono ancora in grado di fare. Ad esempio, un sistema AGI dovrebbe poter applicare principi appresi in fisica per risolvere problemi di economia, linguistica o etica, operando su base astratta e adattiva. Ciò richiede una profonda integrazione tra logiche simboliche, apprendimento statistico e memoria contestuale[139,140].

[139] "The Generality Behind the G: Understanding Artificial General Intelligence (AGI) author-img", scheda a cura di SingularityNET, 6 June 2024, https://singularitynet.io/the-generality-behind-the-g-understanding-artificial-general-intelligence-agi

[140] " What is Artificial General Intelligence (AGI)?, scheda a cura di Digital Ocean, February 19, 2025, https://www.digitalocean.com/resources/articles/artificial-general-intelligence-agi.

Altrettanto critica è la questione dell'allineamento tra i sistemi AGI e i valori umani. Il dibattito sul cosiddetto "*alignment problem*" si concentra sul rischio che agenti dotati di autonomia cognitiva sviluppino comportamenti imprevisti o dannosi, se non progettati con meccanismi di controllo e supervisione adeguati.
Organizzazioni come DeepMind e OpenAI stanno lavorando su approcci come il *Reinforcement Learning from AI Feedback* (RLAIF) e tecniche ibride che integrano l'essere umano nel ciclo decisionale del sistema, per garantire un'evoluzione dell'intelligenza artificiale che rimanga sicura e comprensibile[141,142].

Sebbene una vera AGI sia ancora lontana, si osserva l'emergere di sistemi agentici avanzati, ispirati a quella visione. Questi sistemi combinano grandi modelli linguistici (LLM) con strumenti esterni (API, ambienti di calcolo, sistemi di memoria) e moduli di orchestrazione che definiscono la logica di azione. Si tratta di architetture multilivello articolate in:

- un *model layer* (per comprensione e generazione),

[141] A. DRAGAN, R. SHAH, F. FLYNN, S. LEGG, *Taking a responsible path to AGI*, *Paper* disponibile su Google Deep Mind, 2 April 2025, https://deepmind.google/discover/blog/taking-a-responsible-path-to-agi.

[142] "How we think about safety and alignment", *statement* di OpenAI, https://openai.com/safety/how-we-think-about-safety-alignment.

- un *tool layer* (per operare su dati e ambienti esterni),
- un *orchestration layer* (che governa il ciclo cognitivo percezione-ragionamento-azione).

L'interazione tra questi strati (*layers*) consente al sistema non solo di rispondere a domande, ma di eseguire compiti articolati, pianificare azioni, interagire dinamicamente con l'ambiente e apprendere dai propri errori.

Questi agenti stanno trovando applicazioni in ambito robotico, manifatturiero, decisionale e predittivo, anticipando alcune delle funzionalità attese da una futura AGI.

È proprio lungo questa traiettoria che si muovono i grandi investimenti di attori come Google (con Gemini)[143], Microsoft (con AutoGen e Azure AI Agents) e OpenAI (con il programma Superalignment), nella prospettiva di sviluppare sistemi capaci di eseguire compiti generici in modo sicuro e controllabile[144,145].

[143] S. PICHAI, D. HASSABIS, K. KAVUKCUOGLU, "Introducing Gemini 2.0: our new AI model for the agentic era", *Google Blog*, December 11, 2024, https://blog.google/technology/google-deepmind/google-gemini-ai-update-december-2024.

[144] W. KNIGHT, "Google Reveals Gemini 2, AI Agents, and a Prototype Personal Assistant", *Wired*, December 11, 2024, https://www.wired.com/story/google-gemini-2-ai-assistant-release.

[145] M. WHALIN, "Introducing Azure AI Agent Service", *Azure AI services Blog*, November 19, 2024, https://techcommunity.microsoft.com/blog/azure-ai-services-blog/introducing-azure-ai-agent-service/4298357.

L'Intelligenza Artificiale Generale, dunque, non rappresenta solo un'evoluzione quantitativa della IA odierna, ma una svolta concettuale, potenzialmente paragonabile a quella che, nella storia della specie umana, ha segnato la comparsa del linguaggio simbolico o del pensiero astratto. Per questo motivo, la sua realizzazione dovrà essere accompagnata da una riflessione interdisciplinare sul piano etico, tecnico e istituzionale. La sfida dell'AGI non è solo scientifica: è sociale, culturale e politica.

II.1.5 IA GENERATIVA PER IL LINGUAGGIO: MODELLI DI LINGUAGGIO DI ULTIMA GENERAZIONE (TIPO O1)

I modelli di linguaggio avanzati come GPT-4 e o1 rappresentano un'evoluzione significativa nel campo dell'elaborazione del linguaggio naturale (NLP). Questi modelli combinano enormi quantità di dati con architetture sempre più sofisticate, consentendo di generare testi coerenti, contestualizzati e personalizzati in tempo reale.

Le aziende li stanno adottando per automatizzare la gestione del servizio clienti attraverso chatbot intelligenti, migliorare la

produttività degli impiegati con strumenti di scrittura assistita e ottimizzare i processi di analisi testuale su grandi volumi di dati.

Nel notare le potenzialità degli attuali modelli di GenAI per il linguaggio, è importante ricordarne le limitazioni attuali. Yan LeCun, Vicepresidente e *Chief AI Scientist* di Meta, nel suo intervento al World Economic Forum 2025 di Davos[146] ha commentato: "Penso che la durata dell'attuale paradigma [LLM] sia piuttosto breve, probabilmente da tre a cinque anni", ha detto LeCun. "Penso che entro cinque anni, nessuno sano di mente li userebbe più, almeno non come componente centrale di un sistema di intelligenza artificiale. Penso che [...] assisteremo all'emergere di un nuovo paradigma per le architetture di IA, che potrebbero non avere le limitazioni degli attuali sistemi di intelligenza artificiale".

Queste "limitazioni" inibiscono un comportamento veramente intelligente nelle macchine, afferma LeCun. Ciò è dovuto a quattro motivi chiave: la mancanza di comprensione del mondo fisico; la mancanza di memoria persistente; la mancanza di

[146] L'intervento è su YouTube, qui: https://www.youtube.com/watch?v=MohMBV3cTbg.

ragionamento e la mancanza di capacità di pianificazione complessa. "Gli LLM non sono davvero in grado di fare nulla di tutto questo. - ha detto LeCun - Quindi ci sarà un'altra rivoluzione dell'IA nei prossimi anni. Forse dovremo cambiare il nome, perché probabilmente non sarà generativa nel senso in cui s'intende oggi".

II.1.6 IA PER LA GENERAZIONE DI VIDEO

I modelli di IA per la generazione di video, come VEO o RunwayML, stanno rivoluzionando l'industria audiovisiva. Sono sistemi che consentono di creare contenuti video realistici, partendo da descrizioni testuali o immagini di riferimento, riducendo significativamente tempi e costi di produzione.

Nel settore dell'intrattenimento, tali tecnologie sono utilizzate per creare effetti speciali avanzati[147], generare filmati animati[148] e persino personalizzare contenuti in base alle preferenze degli utenti.

[147] M. EAKIN, "Pixar Used AI to Stoke Elemental's Flame", *Wired*, June 16, 2023, https://www.wired.com/story/pixar-elemental-artificial-intelligence-flames/.

[148] I. FAILES, "The new motion prediction from Wonder Dynamics, and what's coming with facial animation", *Befores & Afters*, April 22, 2025, https://beforesandafters.com/2025/04/22/the-new-motion-prediction-from-wonder-dynamics-and-whats-coming-with-facial-animation/; v. anche "Runway Gen-3 alpha: come creare incredibili video AI da immagini e prompt", *HD Blog*, 23 Settembre 2024, https://www.hdblog.it/tecnologia/articoli/n592613/runway-gen-3-come-creare-video-ai-da-immagini/.

Nell'ambito della formazione e della comunicazione aziendale, la generazione automatizzata di video consente di sviluppare materiali didattici interattivi[149] e spiegazioni visuali su larga scala.

II.2. IA Generativa multimodale: integrazione di testi, immagini e suoni

Con intelligenza artificiale generativa multimodale si fa riferimento ad una tecnologia IA capace di integrare, processare (anche in termini di differenti tipologie di fonti) e generare contenuti in diversi formati, quali testi, immagini, suoni e video, superando i limiti dei modelli (cosiddetti monomodali) che sono in grado di operare su un singolo tipo di dato.

Tale approccio permette di creare o analizzare contenuti complessi e comunicare in modo ancora più ricco e articolato, grazie all'abilità dell'IA multimodale di comprendere e sintetizzare informazioni eterogenee, trasformando input diversi in un unico output integrato. Nel concreto, parole, immagini e suoni possono essere convertiti in rappresentazioni digitali, che vengono analizzate e messe in relazione tra loro, consentendo la

[149] Cfr. per esempio il sito di Colossyan: https://www.colossyan.com/.

creazione di contenuti che vanno ben oltre la semplice produzione di testo, come, ad esempio, generare descrizioni testuali a partire da immagini; produrre sequenze video a partire da brevi sceneggiature, generando contestualmente una colonna sonora abbinata alla narrazione visiva; oppure interpretare il contenuto di un documento fotografato.

Potenziali ambiti di applicazione dell'IA generativa multimodale

L'utilizzo di tecnologie multimodali offre notevoli opportunità di applicazione in vari ambiti e settori di industria, quali per esempio:

Marketing & comunicazione: l'IA multimodale può essere sfruttata per migliorare la qualità della comunicazione verso i propri clienti, generando materiali promozionali o informativi distintivi, che integrano testi, immagini e suoni in maniera omogenea. Inoltre, la capacità di generare tali contenuti complessi in diverse versioni e in tempi estremamente ridotti facilita la possibilità di valutare diverse strategie creative.

Media: l'IA multimodale può essere utilizzata per la produzione di contenuti audiovisivi innovativi (esempi molto noti sono Sora di OpenAI e Veo di Google) o per la generazione di effetti speciali,

l'elaborazione automatizzata di video a partire da una sceneggiatura, la creazione di musica e testi e la realizzazione di storie interattive.

Formazione: la possibilità di presentare concetti anche complessi relativi a pressocché qualsiasi argomento, grazie a una combinazione di formati diversi, favorisce un apprendimento più interattivo sia in ambito aziendale che in ambito formativo educazionale. È possibile, ad esempio, tradurre concetti astratti in rappresentazioni visive e sonore, facilitando la comprensione e spronando la curiosità di dipendenti o studenti, o generare in tempo reale rappresentazioni personalizzate degli argomenti in modo da rispondere alle esigenze di un'*audience* diversificata, contribuendo a rendere la formazione più accessibile e inclusiva.

E-commerce: l'IA multimodale consente di offrire esperienze di acquisto personalizzate, analizzando immagini dei prodotti, recensioni o preferenze degli utenti, ottimizzando la presentazione degli articoli e generando raccomandazioni mirate.

Ottimizzazione dei processi: l'IA multimodale può aiutare ad ottimizzare i processi aziendali e produttivi, consentendo di andare oltre la tradizionale automazione robotica, fondata su una programmazione deterministica delle attività da compiere.

II.2.1 CONSIDERAZIONI RIGUARDO L'ADOZIONE DELL'IA GENERATIVA MULTIMODALE

Nonostante la complessità intrinseca di questi sistemi, l'adozione dell'IA generativa multimodale all'interno, per esempio, di un'azienda non richiede necessariamente ulteriori conoscenze tecniche approfondite. Infatti, da un punto di vista dell'utente finale, l'esperienza e la modalità di interazione con modelli di IA multimodali non si discostano significativamente da quelle con modelli di IA generativa monomodali: l'interfaccia rimane semplice e accessibile, e l'utente può fornire la richiesta (in questo caso con maggiore libertà in termini di formato) ricevendo in risposta i contenuti generati, che possono essere eventualmente salvati per successivi utilizzi.

Small Language Models ed Edge AI

Nel contesto dell'IA generativa, gli *Small Language Models* (modelli linguistici di piccole dimensioni) e l'*Edge AI* rappresentano due tra le principali tendenze tecnologiche, che vedono particolare applicazione nell'ambito di sistemi a basso consumo energetico e per dispositivi mobili e IoT (*Internet of Things*). Uno degli obiettivi degli Small Language Model, infatti,

è quello di consentire l'utilizzo di IA generativa direttamente sui dispositivi locali, abilitando l'elaborazione dei dati in tempo reale e riducendo la dipendenza da infrastrutture centralizzate.

Aspetti innovativi degli Small Language Model

Il concetto di *Small Language Model* fa riferimento a modelli di IA generativa "compatti", che sono caratterizzati dal possedere un "cervello" di dimensioni più ridotte rispetto ai modelli più potenti e che risultano, in conseguenza di ciò, in grado di operare con minori risorse computazionali e, di conseguenza, minori costi.

A differenza quindi dei grandi modelli di intelligenza artificiale generativa che richiedono, nella quasi totalità dei casi, una capacità di calcolo disponibile solo su infrastrutture *cloud*, questi modelli "ridotti" sono ottimizzati per fornire un livello di prestazione soddisfacente in termini di capacità di comprensione e generazione del linguaggio naturale, che non solo si avvicina a quello offerto dai modelli di maggiori dimensioni, ma che ne rende possibile l'utilizzo anche in ambienti con *hardware* e capacità computazionali contenuti quali smartphone, dispositivi di domotica (p.e., *smart speaker*) o altri dispositivi periferici.

Per tal motivo gli Small Language Model rappresentano l'abilitatore ideale per l'*Edge AI* generativa, la quale si pone l'obiettivo primario di permettere la fruizione delle abilità dell'IA generativa direttamente "in loco" sui dispositivi periferici, in alternativa all'invio di dati a *server cloud* remoti per il processamento.

Un tale approccio comporta dunque numerosi vantaggi, quali la riduzione della latenza (elemento critico in contesti dove la rapidità di risposta è fondamentale), maggiore sicurezza dei dati e una maggiore resilienza e autonomia operativa. Grazie all'*Edge AI*, le applicazioni sui dispositivi, infatti, possono reagire più rapidamente a input o stimoli provenienti dal mondo esterno, senza dover fare affidamento su connessioni di rete stabili o costantemente attive, o richiedere un'ampia larghezza di banda.

Principio di funzionamento degli Small Language Model

Come accade per la maggior parte dei Large Language Model (LLM), il funzionamento degli Small Language Model si basa sull'utilizzo di architetture di reti neurali molto avanzate, basate sui cosiddetti "Transformer", con la differenza che vengono utilizzati algoritmi e tecniche di compressione e *fine-tuning ad-*

hoc che consentono di ridurre le dimensioni della rete neurale (e di conseguenza il numero di cosiddetti "parametri", che è il primo indicatore del consumo di risorse del modello) in modo da garantire prestazioni adeguate con risorse computazionali contenute. Gli *Small Language Models* sono dunque "addestrati" per comprendere e generare linguaggio naturale con una struttura interna molto più snella, eliminando componenti superflue che non incidono significativamente sui risultati finali e permettono di ottenere un ottimo compromesso tra qualità e consumi di risorse.

Ulteriori potenzialità degli Small Language Model

L'adozione degli *Small Language Models on Edge* si rivela particolarmente vantaggiosa per la creazione di dispositivi intelligenti che operano in ambienti distribuiti: ad esempio, in un sistema di domotica un dispositivo dotato di questa tecnologia può interpretare comandi vocali e fornire risposte molto velocemente, rendendo più naturale l'interazione con l'utente e senza dover inviare continuamente dati ad un *server* centrale. Inoltre, i dispositivi che eseguono Small Language Model non solo possono operare per periodi più lunghi senza

necessità di ricarica o interventi manutentivi, ma permettono anche di ottenere un miglioramento nella gestione di aspetti di *privacy* e sicurezza dei dati elaborati: infatti, processando le informazioni direttamente sul dispositivo, si riduce il rischio di intercettazioni o attacchi che potrebbero essere compiuti durante il trasferimento dei dati verso le infrastrutture *cloud*, consentendo quindi di rafforzare il controllo sulle informazioni sensibili e rendendo le applicazioni più resilienti rispetto a potenziali minacce esterne.

Graphrag per il recupero efficiente delle informazioni

Il recupero efficiente delle informazioni (in prevalenza in forma testuale) rappresenta ad oggi una delle sfide chiave per sfruttare al meglio le capacità dell'intelligenza artificiale generativa, anche in considerazione del fatto che la quantità di dati potenzialmente oggetto del recupero è in costante crescita e la capacità di estrarre di volta in volta il contenuto più rilevante in tempi rapidi risulta fondamentale. Da questo punto di vista, la *Retrieval-Augmented Generation* (RAG) rappresenta attualmente l'approccio più comunemente impiegato nella realizzazione di applicazioni fondate su IA

generativa in ambito aziendale, ogni qualvolta sia necessario far sì che la generazione di risposte e contenuti possa far leva su informazioni proprietarie.

In gran parte delle applicazioni che utilizzano l'IA generativa, infatti, la RAG è la componente fondamentale per abilitare l'IA generativa ad accedere in maniera rapida ed economica alle informazioni contenute in un *corpus* di documenti aziendali aggiornato e di sicura origine (che, sulla base dell'esperienza, possono essere FAQs, contratti, documenti normativi interni, manuali tecnici, procedure, e-mail, etc.[150]), filtrando e recuperando i passaggi più rilevanti per generare gli output richiesti, andando oltre la conoscenza generica e non certificata del modello fornita dal costruttore durante la fase di addestramento; minimizzando al contempo sia la quantità di dati processati dall'IA che il rischio di "allucinazioni" (ovvero di imprecisioni o errori nelle risposte).

[150] "Che cos'è la *Retrieval-Augmented Generation* (RAG)?", un punto di partenza offerto da Google Cloud, https://cloud.google.com/use-cases/retrieval-augmented-generation?hl=it

Tra gli sviluppi recenti in questo ambito, GraphRAG (*Graph Retrieval-Augmented Generation*) è emerso come un approccio molto promettente che integra, come il nome stesso suggerisce, il concetto di struttura dei grafi[151] nella fase di recupero delle informazioni, con l'obiettivo principale di migliorare significativamente il processo di selezione e recupero.

L'aspetto più innovativo dell'approccio GraphRAG si fonda dunque sull'organizzazione dei dati in forma di grafo, dove ogni nodo rappresenta un'informazione o un concetto e i collegamenti tra i nodi indicano le relazioni e connessioni semantiche esistenti tra le diverse informazioni. È un approccio che consente di andare oltre i sistemi di recupero delle informazioni, che si basano su rappresentazioni basate sulla somiglianza semantica, le quali risultano poco adatte per descrivere il contesto che collega le diverse informazioni e che portano quindi alla generazione di risposte che possono risultare incomplete e poco "spiegabili" agli occhi dell'utilizzatore.

[151] Grafo è in matematica la configurazione formata da un insieme di punti (*vertici* o *nodi*) e un insieme di linee (*archi*) che uniscono coppie di nodi; formalmente è un insieme in cui è definita una relazione di qualunque tipo. Perciò i grafi sono una struttura matematica molto usata nelle applicazioni e si prestano a rappresentare problemi apparentemente molto diversi tra di loro con un linguaggio semplice e unificato.

Uno degli aspetti distintivi di GraphRAG è proprio quindi la capacità di sfruttare le connessioni intrinseche tra le informazioni per migliorare il processo di recupero, grazie al fatto che l'integrazione della struttura a grafo permette di mappare e identificare rapidamente le relazioni chiave e di scartare le informazioni non rilevanti.

Da un punto di vista del funzionamento, GraphRAG si basa su due fasi: nella prima, le informazioni sono pre-processate e organizzate in una apposita struttura a grafo; tale rappresentazione permette di mappare in modo facilmente accessibile le correlazioni tra i vari dati, evidenziando connessioni semantiche che potrebbero sfuggire a un'analisi manuale. Nella seconda - quando viene formulata una richiesta - il sistema utilizza specifici algoritmi di ricerca e di "navigazione" all'interno dell'alberatura del grafo per identificare i nodi e i collegamenti che meglio rispondano a tale richiesta.

Approccio che consente di ridurre il "rumore" presente spesso nei grandi *corpora* documentali, e, concentrandosi sulle connessioni realmente importanti, di ottenere risposte non solo

più rapide, ma soprattutto di qualità significativamente superiore e che riescono a far leva in maniera più efficace su informazioni provenienti da fonti documentali diverse.

Considerazioni per l'adozione

L'adozione dell'approccio GraphRAG abbinato all'IA generativa risulta particolarmente rilevante in ambiti in cui il *corpus* documentale nel quale effettuare la ricerca è notevolmente ampio ed eterogeneo, le relazioni tra le diverse informazioni sono complesse, e la trasparenza (ovvero l'abilità per l'utente di risalire ai contributi primari che hanno guidato la generazione della risposta) è fondamentale. I settori principali dove l'adozione di GraphRAG si sta rivelando particolarmente promettente sono dunque quelli dove la natura stessa dell'attività contribuisce a produrre un patrimonio informativo testuale elevato, come servizi finanziari, telecomunicazioni, utility, ricerca scientifica, media e servizi pubblici.

II.2.2 *LARGE ACTION MODELS* E IPERAUTOMAZIONE

Nel panorama attuale dell'intelligenza artificiale, i Large Action Model (in sigla, LAM), pubblicizzati per la prima volta su larga scala da rabbitAI[152] all'evento CES di Las Vegas nel 2024[153], si stanno delineando come una tecnologia particolarmente promettente, poiché permettono di tradurre in azioni concrete gli output forniti da un'IA generativa, abbinando delle "braccia" al "cervello" costituito dall'IA generativa.

Se infatti, ad esempio, un sistema di IA generativa può facilmente, a fronte di una opportuna richiesta, creare un itinerario di viaggio e sottoporlo all'utente, un LAM è in grado - in totale autonomia e senza alcuna indicazione specifica programmata a priori - di analizzare tale itinerario e "metterlo in pratica", effettuando per esempio le prenotazioni di voli e alberghi, semplicemente navigando su internet o effettuando telefonate, esattamente come farebbe un essere umano.

[152] La rabbitAI, fondata nel 2020, deriva dall'esperienza maturata dai fondatori all'Heidelberg Collaboratory for Image Processing (HCI) https://rabbitai.de/.

[153] R. COSENTINO, "Rabbit R1 presentato al Ces 2024: a metà tra un chatbot e un assistente vocale, nel corpo di uno smartphone", *Login: Corriere della Sera*, 12 gennaio 2024, Link Login: Corriere della Sera.

In un contesto così, questo concetto, che prende il nome di 'iperautomazione', gioca quindi un ruolo sempre più strategico anche in chiave aziendale, poiché può abilitare, ad esempio, l'automatizzazione in maniera pressoché completa di molti processi aziendali complessi, superando di gran lunga le capacità offerte dai sistemi di tradizionale *Robotic Process Automation* e, inoltre, apre le porte all'utilizzo di una robotica sempre più intelligente[154].

Sinergia tra IA e Automazione

Quanto descritto rende quindi evidente che i Large Action Model vanno ben oltre le tradizionali capacità dei Large Language Model, che si limitano a comprendere e generare testi, essendo questi progettati fin dalle basi per orchestrare azioni reali e complesse, interfacciandosi con diversi sistemi informatici e piattaforme *software* proprio come farebbe un dipendente. In pratica, come il nome suggerisce, non si tratta più solo di parlare, scrivere o generare immagini, ma di agire: si possono eseguire comandi, attivare processi e coordinare flussi di lavoro in modo autonomo.

[154] *Technology Vision 2025, AI: A Declaration of Autonomy*, January 2025, *Report* edito da ACCENTURE, //www.accenture.com/us-en/insights/technology/technology-trends-2025.

Risulta quindi chiaro il perché questo nuovo paradigma si sposi perfettamente con l'idea di iperautomazione, un concetto che non si ferma dunque alla semplice automazione di compiti ripetitivi, ma che integra in una maniera molto più potente l'intelligenza artificiale e l'automazione a copertura dell'intero ciclo operativo, dalla raccolta e analisi dei dati, fino alla decisione e all'esecuzione delle azioni con minime necessità di intervento o supervisione umana.

Infatti, il cuore di questi modelli risiede proprio nella loro capacità di "capire" il contesto e tradurre questa comprensione in azioni operative - i LAM sono addestrati a riconoscere condizioni, schemi e pattern all'interno dei dati e quando viene rilevata una determinata situazione, sono in grado di decidere quale azione intraprendere, interagendo direttamente con persone (anche tramite voce sintetica), *software* e piattaforme presenti nell'ecosistema in cui operano fino a raggiungere gli obiettivi richiesti.

Considerazioni per l'adozione

Pur essendo, nel momento in cui si scrive, una tecnologia ancora in fase di maturazione, l'adozione dei Large Action Model abbinata ad una strategia di iperautomazione si prevede offrirà

nel prossimo futuro numerosi benefici, permettendo di ridurre drasticamente i tempi di esecuzione dei processi operativi e automatizzando le attività ripetitive e di *routine* rendendole anche più resilienti, consentendo quindi di focalizzare le proprie risorse umane sui compiti a maggiore valore.

Un altro aspetto sicuramente interessante è la capacità di tali sistemi di adattarsi in tempo reale a modifiche delle condizioni dell'ambiente nel quale interagiscono, individuando rapidamente la soluzione migliore in base alle circostanze.

II.3.1 *AGENTIC AI*

Nella sua forma più fondamentale, un agente di Intelligenza Artificiale Generativa può essere definito come un sistema autonomo progettato per perseguire un obiettivo, osservando l'ambiente circostante e interagendo con esso attraverso gli strumenti a sua disposizione. Gli agenti sono dotati di un elevato grado di autonomia e possono operare indipendentemente dall'intervento umano, in particolare quando vengono forniti obiettivi chiari da raggiungere.

Un aspetto distintivo degli agenti è la loro proattività: non si limitano a rispondere a istruzioni esplicite, ma possono ragionare autonomamente su quale sia la prossima azione più adeguata per avvicinarsi all'obiettivo finale. Questa capacità di auto-direzione consente agli agenti di adattarsi a contesti dinamici e di risolvere problemi complessi anche in assenza di una guida diretta.

Sebbene il concetto di agente in ambito IA sia estremamente ampio e versatile, **l'analisi proposta si concentra su una classe specifica di agenti**, ovvero **quelli generati e gestiti dai modelli di Intelligenza Artificiale Generativa disponibili allo stato attuale della ricerca** e dell'innovazione tecnologica.

Per comprendere il funzionamento interno di un agente, è necessario introdurre i componenti fondamentali che ne determinano il comportamento, le azioni e i processi decisionali. L'insieme di questi elementi può essere descritto come un'architettura cognitiva, una struttura che può essere configurata in molteplici varianti attraverso la combinazione modulare dei diversi componenti. Quindi, dal punto di vista funzionale, **un agente generativo si basa su tre componenti**

essenziali che ne costituiscono il nucleo operativo e che definiscono la sua capacità di percezione, ragionamento e interazione con l'ambiente e ne modellano l'evoluzione nel tempo. In aggiunta, la loro organizzazione all'interno dell'architettura cognitiva dell'agente determina il grado di autonomia, adattabilità e capacità di risoluzione dei problemi che il sistema è in grado di esprimere.

- Il Modello (*Model Layer*): la base cognitiva degli agenti è spesso costituita da modelli di linguaggio (LM), che forniscono capacità di ragionamento e comprensione. Tuttavia, un agente non è solo un modello statico, ma un'entità in grado di interagire attivamente con il mondo esterno.
- Gli Strumenti (*Tool Layer*): i modelli di intelligenza artificiale tradizionali sono vincolati ai dati di addestramento, ma gli agenti possono accedere a strumenti esterni, API e database per ottenere informazioni aggiornate o eseguire azioni nel mondo reale. Ad esempio, un agente può interrogare un sistema di prenotazione per recuperare in tempo reale le informazioni sui voli o eseguire una transazione finanziaria.
- L'Orchestrazione (*Orchestration Layer*): questo livello definisce il ciclo cognitivo dell'agente: raccoglie informazioni,

esegue il ragionamento interno e decide l'azione successiva da intraprendere. A seconda della complessità dell'agente, questo livello può includere semplici regole decisionali o tecniche avanzate come il ragionamento probabilistico e gli algoritmi di apprendimento automatico.

I tre livelli che costituiscono un agente IA interagiscono in modo sinergico per consentire l'esecuzione di compiti complessi. Il modello di linguaggio (*Model Layer*) interpreta l'input dell'utente e genera un'intenzione o un piano d'azione. Questo output viene gestito dal livello di orchestrazione (*Orchestration Layer*), che valuta se l'informazione prodotta è sufficiente o se è necessario consultare una fonte esterna o attivare uno strumento. Se necessario, il *Tool Layer* entra in gioco, consentendo all'agente di interrogare una banca dati, utilizzare un'API o interagire con sistemi esterni.

Una volta ottenuta la risposta dallo strumento, l'orchestratore reintegra l'informazione nel ciclo cognitivo, passando nuovamente il contesto al modello, che può rifinire la risposta o decidere un nuovo passo. Questo ciclo iterativo tra i tre livelli - interpretazione, azione, rielaborazione - è ciò che conferisce agli

agenti la capacità di adattarsi, esplorare alternative e raggiungere un obiettivo, superando i limiti di una semplice interazione domanda-risposta. L'integrazione di questi tre livelli consente agli agenti di agire in modo più sofisticato rispetto ai modelli tradizionali, migliorando la loro capacità di adattarsi dinamicamente alle circostanze e di gestire compiti complessi in ambienti in continua evoluzione.[155]

Nel contesto di un agente, il "modello LLM" si riferisce al modello di linguaggio (*Language Model*, LM) che funge da unità centrale per il processo decisionale dell'agente. Il modello utilizzato da un agente può essere costituito da uno o più modelli di linguaggio di diverse dimensioni (da modelli compatti a modelli di larga scala) e deve essere in grado di eseguire ragionamenti basati su istruzioni e *framework* logici, come ReAct, *Chain-of-Thought* (CoT) o *Tree-of-Thoughts* (ToT)[156].

[155] J. WIESINGER, P. MARLOW, V. VUSKOVIC, *Agents*, White Paper, February 2025, https://www.kaggle.com/whitepaper-agents.

[156] J. WIESINGER, *Agents*, op.cit.

I modelli impiegati possono assumere forme diverse a seconda delle necessità dell'architettura dell'agente:

- Modelli generici, adatti a una vasta gamma di compiti.
- Modelli multimodali, capaci di elaborare input testuali, visivi o sonori.
- Modelli specializzati, sottoposti a un processo di *fine-tuning* per ottimizzarne le prestazioni su specifiche attività.

Un recente articolo di McKinsey[157] prevede che questi sistemi potranno funzionare come colleghi virtuali altamente qualificati, gestendo compiti complessi come la sottoscrizione di prestiti, la documentazione del codice e le campagne di *marketing online*. Sebbene la tecnologia sia ancora nelle fasi iniziali e richieda ulteriori sviluppi tecnici, sta già attirando significativi investimenti da parte di aziende come Google[158], Microsoft[159] e OpenAI, suggerendo che gli agenti potrebbero presto diventare comuni quanto i chatbot odierni.

[157] L. YEE, M. CHUI *ET AL.*, "Why agents are the next frontier of generative AI", *Mckinsey Quarterly*, July 24, 2024, https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/why-agents-are-the-next-frontier-of-generative-ai.

[158] W. KNIGHT, "Google Reveals Gemini 2, AI Agents, and a Prototype Personal Assistant", *Wired*, December 11 2024, https://www.wired.com/story/google-gemini-2-ai-assistant-release/.

[159] N. ROBINS-EARLY, "Microsoft beats Wall Street expectations for fourth quarter in a row amid AI boom", *The Guardian*, 30 April 2025, https://www.theguardian.com/technology/2025/apr/30/microsoft-earnings-report.

II.3.2 SISTEMI MULTI-AGENTI

I sistemi multi-agente (*Multi Agent Systems*, in sigla MAS) rappresentano un'evoluzione rispetto all'intelligenza artificiale tradizionale centralizzata, attraverso delle interazioni locali, i MAS possono risolvere problemi che risultano difficili per i singoli agenti o per sistemi centralizzati, sfruttando la cooperazione per ottenere proprietà globali superiori alla somma delle singole parti.

In effetti, molte architetture multi-agente superano le prestazioni dei singoli agenti, grazie alla diversificazione delle competenze e alla condivisione delle conoscenze, colmando le lacune informative dei singoli modelli.

Questa coordinazione emergente è stata dimostrata in molte simulazioni, come l'esperimento di *hide-and-seek* di OpenAI[160], in cui gli agenti hanno spontaneamente sviluppato strategie complesse e imparato a utilizzare strumenti al di là delle istruzioni ricevute. Tale fenomeno dimostra come i MAS possano sviluppare un "*curriculum* autonomo", in cui la competizione e la cooperazione

[160] Il video "OpenAI Plays Hide and Seek…and Breaks the Game!" è disponibile su You Tube, qui: https://www.youtube.com/watch?v=Lu56xVlZ40M.

generano forme di apprendimento che si auto-migliorano nel tempo, portando a soluzioni inaspettate e innovative.

Un'ulteriore sfida è data dalla difficoltà di coordinare agenti eterogenei e garantire il loro allineamento con obiettivi globali e valori umani. L'efficacia dei MAS dipende dalla loro capacità di lavorare in sinergia, ma tale coordinazione non è garantita a priori. Se gli agenti interpretano male i segnali provenienti dagli altri, utilizzano funzioni di ricompensa mal progettate o privilegiano obiettivi locali rispetto alla finalità del sistema, il risultato può essere una frammentazione del MAS in sottogruppi competitivi, oppure l'ingresso in modalità di fallimento sistemico.

Per evitare questi rischi, è necessario sviluppare solidi *framework* di *governance*, che definiscano regole chiare per il comportamento e l'interazione degli agenti. La *governance*, in questo contesto, si articola su due livelli:

- Protocolli tecnici: meccanismi per risoluzione dei conflitti, sistemi di *fail-safe* per prevenire decisioni pericolose, e meccanismi di *override* per consentire agli esseri umani di intervenire nei processi decisionali.

- Normative organizzative: politiche che definiscono limiti etici, criteri di trasparenza e regole per la supervisione delle decisioni prese dagli agenti autonomi.

BOX 1. *ECONOMY OF TRUST* E *TRUSTED NETWORKS*. *BLOCKCHAIN* E IA: INTERAZIONE E SFIDE, FILIGRANATURA E TRACCIABILITÀ DI DATI.

Oggi la fiducia rappresenta nell'economia digitale un pilastro fondamentale del valore d'impresa. Le grandi aziende operano sempre più in una "economia della fiducia", in cui integrità e autenticità dei dati sono aspetti fondamentali. Tecnologie emergenti come la *blockchain* e l'intelligenza artificiale offrono capacità complementari per rafforzare tale fiducia: la *blockchain* garantisce l'inalterabilità e la tracciabilità dei dati attraverso registri immutabili, mentre l'IA consente di verificare e filigranare le informazioni. La combinazione di queste tecnologie dà vita a reti affidabili che permettono alle imprese di validare con certezza dati, identità e transazioni.

Nel settore finanziario, i *deepfake* sono stati utilizzati per esempio per imitare la voce di CEO[161], ingannando dipendenti e inducendoli ad autorizzare bonifici fraudolenti. Un *report* ha evidenziato che gli incidenti legati ai *deepfake* nel settore *fintech* sono aumentati del 700% nel 2023, a causa della crescente accessibilità di strumenti IA per la generazione di contenuti falsificati[162].

In parallelo, la frode con identità sintetiche - in cui criminali utilizzano l'IA per creare persone fittizie con credenziali apparentemente valide - sta divenendo il crimine finanziario a crescita più rapida. Si stima che solo nel

[161] Cfr. https://www.businessinsider.com/bank-account-scam-deepfakes-ai-voice-generator-crime-fraud-2025-5.

[162] Sulle frodi bancarie, cfr. https://www.deloitte.com/us/en/insights/industry/financial-services/deepfake-banking-fraud-risk-on-the-rise.html.

2023 abbia causato perdite per 35 miliardi di dollari[163]. Fenomeni che non minacciano solo la sicurezza economica, ma erodono la fiducia nei sistemi digitali, rendendo sempre più necessaria l'adozione di strategie avanzate di verifica dell'identità[164].

La *blockchain* è spesso descritta come una "macchina della fiducia[165]", grazie alla sua capacità di mantenere un registro immutabile dei dati. Una volta registrata, un'informazione è estremamente difficile da modificare, garantendo un elevato livello di integrità. Ad esempio, attraverso la marcatura temporale e la 'notarizzazione' su *blockchain*, le aziende possono stabilire una catena di custodia verificabile, dalla creazione dei dati alla loro archiviazione. Ciò assicura che qualsiasi documento o *record* (contratti, immagini, *log* di transazioni, ecc.) possa essere successivamente verificato come versione originale e inalterata.

Nei settori della *supply chain*[166] e dell'*audit*[167], tale tracciabilità aumenta la fiducia nella genuinità dei documenti. Si evidenzia che la tracciabilità è stata scelta dalla stessa Commissione per l'Intelligenza Artificiale del dipartimento per l'Informazione e l'Editoria della Presidenza del Consiglio dei Ministri, che ha garantito l'integrità informativa e la marcatura temporale della relazione[168] dalla stessa elaborata, poi consegnata alla Presidenza del Consiglio, proprio

[163] THE FEDERAL RESERVE, scheda a cura della FedPayments Improvement, "When Synthetic Identities and Artificial Intelligence Collide", https://fedpaymentsimprovement.org/news/blog/artificial-intelligence-insights-added-to-synthetic-identity-fraud-mitigation-toolkit.

[164] Cfr. https://www.zyphe.com/blog/what-is-zero-knowledge-proof-in-kyc-verification.

[165] Cfr. https://www.ilsole24ore.com/art/blockchain-macchina-fiducia-o-strumento-controllo-ABKSNosB.

[166] Cfr. https://hbr.org/2020/05/building-a-transparent-supply-chain.

[167] Cfr. https://www.agendadigitale.eu/documenti/come-blockchain-e-ai-trasformeranno-gli-audit/.

[168] Cfr. https://agenparl.eu/2024/03/26/intelligenza-artificiale-barachini-relazione-certificata-con-marcatura-temporale-blockchain/.

utilizzando la tecnologia *blockchain*, anche grazie allo sviluppo di uno *smart-contract* apposito, redatto con la collaborazione dell'Osservatorio *Blockchain & Web3* del Politecnico di Milano.

Con il rafforzamento delle difese contro le frodi digitali, emerge una nuova frontiera all'intersezione tra Web3 (tecnologie decentralizzate) e IA: la verifica dell'identità umana nel mondo digitale. In un'epoca in cui l'IA può creare identità false, diventa fondamentale provare che una persona con cui si interagisce *online* sia reale. Due innovazioni offrono una possibile soluzione:

- *Zero-Knowledge Proofs*[169] (ZKPs) per la verifica dell'identità senza rivelare dati sensibili.
- Sistemi KYC[170] (*Know Your Customer*) di nuova generazione, che combinano la garanzia di fiducia della *blockchain* con la capacità dell'IA di validare le identità.

Allo stesso modo, le identità *self-sovereign identity* (SSI) consentono al singolo di gestire le proprie credenziali verificate in un portafoglio digitale, presentando prove crittografiche solo quando necessario.

L'IA facilita questo processo automatizzando la verifica iniziale (scansione documenti, confronto con *database*, riconoscimento facciale e verifica di *liveness*), per poi valutare le prove presentate nelle future autenticazioni. Un esempio teorico: una banca potrebbe accettare una ZKP che dimostri la verifica dell'identità da parte di un'autorità di fiducia, richiedendo comunque un rapido controllo facciale per confrontare il volto dell'utente con la credenziale biometrica. Tutto ciò avviene in pochi secondi su uno *smartphone*, senza necessità di trasmettere dati sensibili.[171]

[169] Cfr. https://en.wikipedia.org/wiki/Zero-knowledge_proof.

[170] Cfr. https://kpmg.com/ky/en/home/insights/2018/02/blockchain-kyc-utility-fs.html.

[171] Art. cit, v. nota 164.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

III.

LA SOSTENIBILITÀ

Con il contributo di

Pietro Biancone, Davide Calandra

III.4.1 DIGITALIZZAZIONE E INTELLIGENZA ARTIFICIALE. SI, MA QUALE È L'IMPATTO AMBIENTALE?

La digitalizzazione e l'intelligenza artificiale stanno rivoluzionando il nostro quotidiano e il mondo del lavoro, ma è fondamentale non trascurare le implicazioni ambientali di tali tecnologie. In questo contesto, due approcci distinti ma

complementari tra loro stanno emergendo con forza: intelligenza artificiale sostenibile e intelligenza artificiale per la sostenibilità[172]. Il primo si concentra sulla riduzione dell'impatto ambientale delle tecnologie di IA lungo tutto il loro ciclo di vita, mentre il secondo mira a impiegare l'IA come strumento per affrontare direttamente le sfide ambientali e sociali.

In un recente contributo pubblicato sulla rivista scientifica *Journal of Cleaner Production*[173] si evidenzia come l'addestramento di modelli di IA, per esempio GPT-3, con i suoi 175 miliardi di parametri, abbia consumato circa 1.287 MWh di elettricità durante la fase di *training*, producendo all'incirca 552 tonnellate di CO_2 equivalenti: un valore comparabile alle emissioni annuali di 123 automobili alimentate a benzina (ALZOUBI E MISHRA, 2024). Tuttavia, modelli più recenti come GPT-4 (2023) e LLaMA 3.1 405B (2024) hanno raggiunto rispettivamente 5.184 e 8.930

[172] KPMG (2023), "Decoding Sustainable AI vs. AI AI for Sustainability, Differences and Impacts", testo disponibile al link.

[173] ALZOUBI, Y. I., MISHRA, A., "Green artificial intelligence initiatives: Potentials and challenges", *Journal of Cleaner Production*, Vol. 468, 25 August 2024, 143090, https://doi.org/10.1016/j.jclepro.2024.143090.

tonnellate di CO_2, evidenziando una tendenza crescente nell'impatto ambientale dovuto ai processi di addestramento[174].

In questo panorama, il progetto cinese DeepSeek ha lanciato una sfida alla "corsa ai parametri" dei LLM. La release DeepSeek-V3 (febbraio 2025) impiega un'architettura *Mixture-of-Experts* con 671 miliardi di parametri totali, ma soltanto 37 miliardi attivi per *token*, contenendo l'addestramento in circa 2,8 milioni di ore-GPU H800 e in un *budget* inferiore ai sei milioni di dollari, a parità di *performance* con modelli assai più energivori[175]. Il caso DeepSeek dimostra che ridimensionare i modelli senza sacrificare le prestazioni è possibile e apre la strada a una "IA frugale" con minore impronta carbonica.

Tale consumo non è limitato solo alla fase di *training*; i data center - infrastrutture essenziali per il funzionamento di questi sistemi - rappresentano attualmente tra il 2 e il 4% del consumo energetico globale in mercati come gli Stati Uniti, la Cina e l'Unione Europea[176]. Inoltre, l'energia impiegata per alimentare

[174] AI Index Report (2025), *Artificial Intelligence Index Report*, p. 73, disponibile al link.

[175] RYAN A. in ApX (2025), "GPU Requirements Guide for DeepSeek Models (V3, All Variants)", January 18, 2025, https://apxml.com/posts/system-requirements-deepseek-models.

[176] IEA (2024), "What the data centre and AI boom could mean for the energy sector", disponibile al link.

tali infrastrutture è spesso di origine non rinnovabile, esacerbando l'impatto in termini di emissioni di gas serra.

Un aspetto critico riguarda inoltre il raffreddamento dei server. Secondo un recente articolo apparso su *Forbes*[177], il raffreddamento dei server IA consuma molta acqua, con i data center che utilizzano torri di raffreddamento e meccanismi ad aria per dissipare il calore, causando l'evaporazione di fino a nove litri di acqua per kWh di energia utilizzata. Altri studi, come il rapporto sull'utilizzo energetico dei data center degli Stati Uniti pubblicato dal Laboratorio Nazionale Lawrence Berkeley[178], mostrano un incremento significativo del consumo idrico. Se nel 2014, il consumo si assestava a 21,2 miliardi di litri d'acqua, nel 2023 la cifra è salita a 66 miliardi di litri totali, con un incremento del 211% in meno di dieci anni.

L'impronta idrica dell'intelligenza artificiale (IA) varia in modo significativo a seconda di dove essa è addestrata e ospitata. Ad

[177] FORBES (2024), "AI Is Accelerating the Loss of Our Scarcest Natural Resource: Water", February 25, 2024, disponibile al link.

[178] SHEHABI A., SMITH S., SARTOR D., BROWN R., HERRLIN M. (2016), "United states data center energy usage report", Ernest Orlando Lawrence Berkeley National Laboratory, p.28, disponibile al link.

esempio, l'IA consuma tra 1,8 e 12 litri di acqua per ogni kWh di energia utilizzata nei data center globali di Microsoft, con Irlanda e stato di Washington rispettivamente come luoghi più e meno efficienti dal punto di vista idrico. Questa variabilità dipende non solo dalla tecnologia di raffreddamento scelta, ma anche dal mix energetico della rete locale, con impatti maggiori in aree dove l'energia proviene da fonti fossili ad alta intensità idrica[179].

Il secondo spunto invita a "non dare per scontato" l'interazione con l'IA. In un post su X del 18 aprile 2025 il CEO di OpenAI, Sam Altman, ha dichiarato che l'aggiunta di formule di cortesia nei prompt - «*please*», «*thank you*» - costa all'azienda «decine di milioni di dollari» l'anno in elettricità e raffreddamento perché ogni parola extra genera token da processare[180]. Un successivo articolo del *New York Times* ha ripreso l'episodio per evidenziare i consumi invisibili degli LLM[181]. L'esempio, pur aneddotico, mostra come scelte comunicative apparentemente minime possano amplificare

[179] GARG A., KITSARA I., BÉRUBÉ S. (2025), "The hidden cost of AI: Unpacking its energy and water footprint", *OECD.AI*, February 26, 2025, disponibile al link.

[180] Post di Sam Altman su X, disponibile al link.

[181] THE NEW YORK TIMES (2025), "Saying 'Thank You' to ChatGPT Is Costly. But Maybe It's Worth the Price", April 24 2025, disponibile al link.

l'impronta ambientale degli LLM e suggerisce di progettare interfacce che riducano l'*overhead* testuale superfluo.

Nel quadro della sostenibilità, è fondamentale inquadrare l'IA anche attraverso la lente ESG (*Environmental, Social, Governance*). Sul piano ambientale, significa promuovere lo sviluppo di modelli più efficienti dal punto di vista energetico, l'uso di materiali sostenibili nei componenti *hardware* e la riduzione dei rifiuti elettronici. Da una prospettiva sociale, l'attenzione si concentra sull'uso etico dell'IA, il rispetto dei diritti umani, la protezione dei dati personali e l'inclusività. Infine, il pilastro della *governance* richiede trasparenza, *accountability* e strutture etiche per la supervisione dello sviluppo e impiego dei sistemi IA[182].

Il dibattito sull'impatto ambientale dell'IA si inserisce in una tradizione di studi che considerano il cosiddetto "*rebound effect*"[183] ovvero il fenomeno per cui miglioramenti in efficienza non sempre comportano una riduzione della domanda, ma

[182] KPMG (2023), "Decoding Sustainable AI vs. AI AI for Sustainability", disponibile al link.

[183] ERTEL, W., & BONENBERGER, C. (2025), "Rebound Effects Caused by Artificial Intelligence and Automation in Private Life and Industry", *Sustainability*, 17(5), p. 1988 ss., https://doi.org/10.3390/su17051988.

possono addirittura portare a un aumento del consumo complessivo di risorse[184]. Un esempio chiaro è rappresentato dai modelli di IA ottimizzati per il risparmio energetico: sebbene consumino meno energia per unità di calcolo, il loro crescente utilizzo su larga scala potrebbe comunque incrementare il fabbisogno complessivo di risorse e infrastrutture informatiche.

In questo contesto, il modello dei tre ordini di effetto proposto da Hilty e Hercheui nel 2010[185] offre una chiave interpretativa fondamentale: il primo ordine riguarda gli impatti diretti lungo il ciclo di vita del prodotto (produzione, uso e smaltimento). Ad esempio, un *chip* IA richiede l'estrazione di minerali rari come il litio e il cobalto, con impatti ambientali significativi già nella fase produttiva. Il secondo ordine evidenzia i benefici ambientali derivanti da miglioramenti di efficienza e ottimizzazione dei processi. Un esempio concreto è l'uso dell'IA per ottimizzare i consumi di energia negli edifici intelligenti, riducendo gli

[184] WILLENBACHER, M., HORNAUER, T., & WOHLGEMUTH, V. (2021), "Rebound effects in methods of artificial intelligence", in *Advances and New Trends in Environmental Informatics*, Springer International Publishing, Cham, pp. 73-85, https://doi.org/10.1007/978-3-030-88063-7_5.

[185] HILTY, L. M., & HERCHEUI, M. D. (2010), "ICT and sustainable development". in *IFIP International Conference on Human Choice and Computers*, Springer Berlin Heidelberg, pp. 227-235, https://doi.org/10.1007/978-3-642-15479-9_22.

sprechi. Infine, il terzo ordine, coglie gli impatti negativi generati dal *rebound effect*, dove un aumento della produzione può, in ultima analisi, incrementare il consumo di energia e risorse. Ad esempio, se un'azienda riduce il consumo di energia per unità di calcolo grazie all'IA, ma aumenta la produzione di servizi IA, il consumo totale potrebbe comunque crescere.

Ulteriori analisi basate su dati panel globali[186] suggeriscono che lo sviluppo dell'IA possa contribuire a ridurre le impronte ecologiche e le emissioni di carbonio, soprattutto quando applicata all'efficientamento energetico e alla gestione delle reti elettriche. Tuttavia, tali benefici possono variare a seconda delle politiche energetiche e delle infrastrutture dei singoli paesi.

Di fronte a tali sfide, le aziende tecnologiche stanno adottando un ventaglio articolato di strategie per contenere l'impatto ambientale delle infrastrutture e dei modelli di intelligenza artificiale.

[186] CHENG, Y., ZHANG, Y., WANG, J., & JIANG, J. (2023), "The impact of the urban digital economy on China's carbon intensity: spatial spillover and mediating effect", *Resources, Conservation and Recycling*, Vol. 189, 2023, p. 106762, ss., https://doi.org/10.1016/j.resconrec.2022.106762.

Una delle azioni più incisive riguarda l'approvvigionamento energetico. Ad esempio, molte imprese leader, come Google, Microsoft e Amazon[187], stanno stipulando contratti per l'acquisto diretto di energia rinnovabile e investendo in impianti solari, eolici e, più di recente, in reattori nucleari modulari per alimentare i propri data center IA con fonti a basse emissioni. Ciò permette una riduzione sostanziale delle emissioni legate al funzionamento dei server.

Sul fronte dell'efficienza computazionale, le imprese stanno tentando lo sviluppo di modelli IA meno energivori, attraverso tecniche come il *model pruning*, *quantization*, *distillation* ed *edge computing*, che permettono di ottenere prestazioni comparabili a costi energetici inferiori. L'uso dell'*edge computing* consente inoltre di ridurre la dipendenza dai data center centralizzati, migliorando l'efficienza ambientale complessiva[188].

[187] AGENDADIGITALE.EU (2024), "La corsa al nucleare delle *Big Tech*: alleanze strategiche per l'IA", 4 novembre 2024, disponibile al link.

[188] MANCHIKANTI D.R. (2025), "Digital Waste Mitigation in AI and Cloud Computing: A Comprehensive Framework for Environmental Sustainability", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, Vol. 11, No. 1, January-February 2025, https://doi.org/10.32628/CSEIT251112147.

Anche la gestione del calore residuo rientra tra le pratiche innovative: alcune aziende stanno recuperando il calore generato dai server per riscaldare edifici vicini o alimentare reti di teleriscaldamento, trasformando uno scarto energetico in risorsa utile[189]. Infine, si sta consolidando una maggiore attenzione alla trasparenza dei consumi, con l'integrazione di metriche ambientali nei bilanci di sostenibilità. L'utilizzo di modelli IA per la contabilità ambientale consente di migliorare l'accuratezza e la tempestività della rendicontazione, facilitando il monitoraggio delle emissioni e promuovendo una cultura della responsabilità ambientale all'interno delle organizzazioni[190].

III.4.2 Il bilancio ambientale all'epoca dell'intelligenza artificiale

L'evoluzione delle norme e degli standard di rendicontazione ambientale sta ridefinendo in modo profondo il concetto di bilancio

[189] Greenpeace (2024), *Clean Cloud 2024 - Tracking Renewable Energy Use in China's Tech Industry*, July 2024, disponibile al link.

[190] World Economic Forum. (2025), *Artificial Intelligence's Energy Paradox: Balancing Challenges and Opportunities*, White paper, Geneva, January 2025, disponibile al link.

ambientale, trasformandolo da strumento di mera rendicontazione a leva strategica per la *governance* e la competitività.

A livello europeo, la svolta più recente è rappresentata dalla Direttiva (UE) 2022/2464, nota come *Corporate Sustainability Reporting Directive* (CSRD), che modifica la precedente Direttiva 2013/34/UE, introducendo obblighi più stringenti in materia di comunicazione societaria sulla sostenibilità. A questa cornice normativa si affiancano i nuovi *European Sustainability Reporting Standards* (ESRS), elaborati dall'EFRAG (*European Financial Reporting Advisory Group*) con l'obiettivo di fornire un quadro tecnico e metodologico unificato per il *reporting* di sostenibilità, incluso quello ambientale.

L'approccio sancito da CSRD e ESRS si basa sul principio della cosiddetta "*double materiality*", secondo cui le imprese sono tenute a rendicontare non solo come i fattori ambientali, sociali e di *governance* (ESG) influiscano sul proprio modello di *business*, ma anche come le attività aziendali abbiano effetto sull'ambiente e sulla società.

In particolare, si richiede una valutazione integrata di più dimensioni, tra cui emissioni di gas serra (*Scope* 1, 2 e 3), consumo e qualità delle risorse naturali, produzione di rifiuti ed effetti sulla biodiversità. L'attenzione al ciclo di vita dei prodotti (*Life Cycle Assessment*) e alla completa catena del valore (*supply chain*) evidenzia come l'impegno per la sostenibilità non possa prescindere da un intervento coordinato su tutti i processi aziendali.

In questo contesto, il bilancio ambientale assume una rilevanza cruciale anche per le aziende leader nel settore digitale, in cui l'adozione di tecnologie di intelligenza artificiale (IA) comporta un impiego intensivo di infrastrutture IT e data center. L'alta potenza di calcolo richiesta per gestire algoritmi avanzati può tradursi in consumi energetici ingenti, con un conseguente impatto significativo sul fronte delle emissioni di gas serra. La standardizzazione proposta dagli ESRS, improntata a indicatori quantitativi e a criteri di comparabilità e verificabilità, spinge le imprese a misurare in modo sempre più accurato i propri consumi, a identificare le aree di inefficienza e a definire obiettivi di riduzione delle emissioni in linea con i target internazionali di contrasto al cambiamento climatico (ad es., quelli proposti dall'Accordo di Parigi).

Alcuni casi, p.e. Apple e Samsung, dimostrano come i grandi provider tecnologici stiano progressivamente adeguandosi alle nuove richieste di trasparenza e responsabilità. Secondo l'*Environmental Progress Report 2024*[191], Apple ha certificato i primi prodotti a zero emissioni di carbonio, integrando l'uso di fonti rinnovabili al 100% nei propri data center e realizzando un piano di riduzione e compensazione delle emissioni lungo la filiera (incluso il coinvolgimento dei fornitori, decisivo per affrontare lo *Scope* 3). Samsung, attraverso il *Sustainability Report 2024*[192], mostra come l'obiettivo *net zero* per *Scope* 1 e 2 entro il 2030 sia perseguito mediante la conversione quasi totale dei siti produttivi all'energia rinnovabile e l'introduzione di semiconduttori a basso consumo energetico, capaci di ridurre del 30% i consumi rispetto ai modelli precedenti.

All'interno dei framework europei, il bilancio ambientale va ben oltre un mero prospetto delle emissioni: diventa uno strumento di programmazione, gestione e coinvolgimento degli stakeholder, finalizzato a migliorare le performance ambientali

[191] APPLE, *2024 Environmental Progress Report*, covering fiscal year 2023, disponibile al link.

[192] SAMSUNG, *Samsung Electronics Releases 2024 Sustainability Report*, June 2024, disponibile al link.

in tutte le fasi del ciclo produttivo. L'impostazione seguita dall'EFRAG insiste sull'obbligo di integrare i dati ambientali (inclusi i parametri relativi all'IA e ai consumi dei data center) nei bilanci e nelle relazioni sulla gestione, affinché vengano valutati assieme agli indicatori finanziari per delineare un quadro veritiero e completo dell'andamento aziendale. Questo passaggio rappresenta una rivoluzione culturale, poiché spinge le imprese a considerare i fattori ESG non più come un *addendum* volontario, ma come variabili di performance essenziali per la sopravvivenza e lo sviluppo del *business*.

Nel settore dell'IA, in particolare, l'adozione di un *environmental accounting* rigoroso implica la misurazione accurata dei consumi associati all'addestramento degli algoritmi (che spesso richiedono un uso massivo di server ad alte prestazioni), l'ottimizzazione dei sistemi di raffreddamento e la progettazione di modelli computazionali meno energivori. Ciò comporta un impegno congiunto tra reparti di ricerca e sviluppo, responsabili della sostenibilità ed esperti di *supply chain management*, in un'ottica di innovazione che coniughi

efficienza operativa e riduzione dell'impronta carbonica. Le linee guida degli ESRS, inoltre, enfatizzano la necessità di un riesame periodico degli investimenti e dei processi per favorire l'adozione di soluzioni circolari, come il recupero di materiali dai dispositivi dismessi o la promozione di economie di scala virtuose in grado di abbattere i costi ambientali legati alla logistica.

Tuttavia, l'espansione delle pratiche di rendicontazione ambientale nel campo dell'intelligenza artificiale espone anche al rischio di fenomeni di *greenwashing* o *bluewashing*, in cui imprese e sviluppatori enfatizzano in modo opportunistico l'etica degli algoritmi o la sostenibilità dei modelli senza un'effettiva coerenza tra comunicazione e pratiche operative.

La retorica della "IA etica" o "neutrale" può così mascherare impatti significativi in termini di consumi energetici, uso di risorse e disuguaglianze algoritmiche, minando la credibilità delle metriche ESG e compromettendo la fiducia degli

stakeholder[193]. Per questo motivo, le autorità regolatorie e gli organismi di standardizzazione insistono sempre più sulla necessità di *audit* ambientali indipendenti, verifiche di terza parte e controlli trasversali sulle dichiarazioni ambientali delle imprese *tech*[194].

Le logiche introdotte dalla CSRD e dagli ESRS promuovono infine la trasparenza e l'affidabilità dei dati ambientali, prescrivendo l'impiego di metodologie di calcolo uniformi e la validazione per conto di terze parti (*assurance*). Tale requisito incentiva le grandi imprese a dotarsi di sistemi di reportistica evoluti che consentano di tracciare, aggregare e rendere pubbliche le informazioni rilevanti su emissioni, consumi e impatti correlati. Ne deriva un cambiamento sostanziale nella relazione con gli investitori, sempre più interessati a valutare la capacità dell'azienda di gestire i rischi climatici e di cogliere le opportunità di un mercato in transizione verso la neutralità

[193] SIMION, R. (2024), "Eco-Frauds: The Ethics and Impact of Corporate Greenwashing", *Studia Universitatis Babeș-Bolyai Philosophia*, Vol. 69(2), pp. 7-26, https://doi.org/10.24193/subbphil.2024.2.01.

[194] DE SILVA LOKUWADUGE, C. S. & DE SILVA, K. M., (2022) "ESG Risk Disclosure and the Risk of Green Washing", *Australasian Accounting, Business and Finance Journal*, Vol. 16(1), August 2022, pp. 146-159, https://doi.org/10.14453/aabfj.v16i1.10.

carbonica. In parallelo, le amministrazioni pubbliche e la società civile dispongono di strumenti più efficaci per monitorare e premiare i comportamenti virtuosi, alimentando un circolo virtuoso in cui la *performance* ambientale diventa un indice di maturità e affidabilità dell'organizzazione.

Alla luce di questo scenario normativo e di mercato, il bilancio ambientale assume una valenza strategica, non più relegata a un esercizio di conformità formale. Grazie ai nuovi standard di *reporting*, i temi di sostenibilità si intrecciano con la programmazione industriale, spingendo le imprese a sviluppare modelli di *business* resilienti e a sperimentare soluzioni tecnologiche avanzate, come l'intelligenza artificiale a basso consumo e l'ottimizzazione "*green*" dei data center.

Coniugare IA, sviluppo tecnologico e responsabilità ambientale rappresenta dunque una priorità, in cui la rendicontazione completa funge da catalizzatore per i processi di miglioramento continuo e per un dialogo costruttivo con l'ecosistema degli stakeholder. In questo modo, le aziende capaci

di integrare le logiche ESG nei propri sistemi decisionali possono non solo rispondere alle richieste normative, ma anche generare valore duraturo, contribuendo alla transizione verso un'economia a ridotto impatto di carbonio.

III.3 Infrastrutture e intelligenza artificiale: ottimizzazione energetica e costi

L'elevato impatto energetico dei sistemi di intelligenza artificiale non dipende soltanto dai processi di addestramento e inferenza degli algoritmi, ma anche dalle infrastrutture fisiche che li supportano, in particolare i data center. Questi impianti richiedono una gestione attenta dei consumi elettrici e idrici, nonché ingenti investimenti per il raffreddamento e la sicurezza operativa.

Secondo i dati dell'International Energy Agency (2024)[195], nelle grandi economie come gli Stati Uniti, la Cina e l'Unione Europea, i data center rappresentano tra il 2% e il 4% del consumo energetico globale, con prospettive di crescita significativa entro il 2030. Di fronte a questi numeri, l'ottimizzazione

[195] IEA (2024), "What the data centre and AI boom could mean for the energy sector", disponibile al link.

infrastrutturale diventa un elemento chiave per coniugare sviluppo tecnologico e sostenibilità, soprattutto in quelle economie dove i data center e l'economia di internet rappresentano uno dei motori dominanti della crescita economica (p.e., in Irlanda).

Dunque, l'adozione di soluzioni IA per l'autogestione energetica dei data center rappresenta un'evoluzione significativa verso la sostenibilità. Aziende come Google e Microsoft utilizzano algoritmi di *machine learning* per regolare automaticamente la ventilazione, la distribuzione del carico di lavoro e l'utilizzo dell'energia in base ai flussi operativi, riducendo il consumo energetico fino al 40% (p.e., Google DeepMind, 2016[196]). Tali soluzioni, che rientrano nell'ambito dell'intelligenza artificiale sostenibile, abbattono i costi operativi e permettono di limitare l'uso di risorse idriche attraverso sistemi di raffreddamento ottimizzati.

Anche la scelta della localizzazione geografica dei data center influisce sensibilmente sui costi e sull'efficienza ambientale. Infrastrutture situate in aree con clima freddo, fonti rinnovabili

[196] GOOGLE DEEP MIND (2016), "DeepMind AI Reduces Google Data Centre Cooling Bill by 40%", 20 July 2016, disponibile al link.

disponibili e accesso abbondante all'acqua, come la Scandinavia o il Canada, permettono di ridurre la domanda di energia per il raffreddamento e di massimizzare l'impiego di elettricità a basso impatto carbonico[197]. Questo principio è alla base della strategia adottata da numerosi operatori del settore per delocalizzare le proprie infrastrutture in regioni più sostenibili.

L'integrazione tra intelligenza artificiale e sistemi energetici intelligenti (*smart grids*) apre inoltre scenari promettenti per la gestione dei picchi di domanda, la previsione dei consumi e il bilanciamento della rete elettrica. Attraverso tecnologie predittive, è possibile distribuire le operazioni di calcolo nei momenti di minor costo energetico, riducendo le emissioni e ottimizzando il bilancio economico[198]. In questo modo, l'IA non è solo una fonte di consumo, ma anche uno strumento per gestire in modo intelligente le risorse.

[197] DEPOORTER, V., ORÓ, E., & SALOM, J, (2015), "The location as an energy efficiency and renewable energy supply measure for data centres in Europe", *Applied Energy*, 2015, vol. 140, pp. 338-349, https://doi.org/10.1016/j.apenergy.2014.11.067

[198] HABBAK, H., MAHMOUD, M., METWALLY, K., FOUDA, M. M., & IBRAHEM, M. I., (2023), "Load forecasting techniques and their applications in smart grids", *Energies*, 2023, vol. 16(3), p. 1480 ss. https://doi.org/10.3390/en16031480.

Infine, la crescente attenzione ai modelli IA a bassa intensità computazionale - come gli algoritmi "distillati", i modelli quantizzati o *edge computing* - offre un'alternativa concreta per contenere sia l'impatto ambientale che i costi operativi. Sono soluzioni che richiedono meno potenza di calcolo e possono essere eseguite localmente su dispositivi con capacità limitate, riducendo la dipendenza dai data center centralizzati. **Pertanto, l'ottimizzazione energetica delle infrastrutture legate all'IA rappresenta una leva strategica fondamentale** per ridurre le esternalità ambientali, abbattere i costi e allinearsi agli obiettivi ESG. Solo attraverso un approccio integrato - che combini soluzioni tecnologiche avanzate, pianificazione logistica e responsabilità ambientale - sarà possibile rendere le infrastrutture digitali sostenibili nel lungo periodo.

Diventa urgente una riflessione sistemica sull'impatto energetico dell'IA: come suggerito da Contucci (2024), è necessario sviluppare una vera e propria 'termodinamica dell'apprendimento automatico' per evitare che i costi energetici compromettano gli obiettivi climatici[199].

[199] P. CONTUCCI, (2024), "Intelligenza Artificiale: rischio energetico e climatico", *Rivista Il Mulino*, 2 luglio 2024, https://www.rivistailmulino.it/a/ia-rischio-energetico-e-climatico.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

IV.

LE SFIDE ETICO-SOCIALI

Con il contributo di

Lorenzo Maternini, Marco Tasca, Rosanna Spanò,

Alessandro Massolo, Michele Gerace

IV.1. LA SFIDA DELLA TRASPARENZA E DELLA RESPONSABILITÀ NEGLI ALGORITMI DI IA [RS1]

Il termine *accountability* è stato utilizzato in molti modi diversi e applicato in una vasta gamma di contesti[200]. Spesso, è stato

[200] SPANÒ, R. (2016), *L'evoluzione dei sistemi di Management Accounting nelle aziende sanitarie. Accountability e fattori di complessità*, G. Giappichelli Editore Torino.

adoperato come un termine suggestivo per rafforzare argomentazioni fragili, evocare immagini di affidabilità, correttezza e giustizia o per rispondere a critiche. Tuttavia, ciò ha portato al rischio che il concetto perda valore analitico, trasformandosi in un contenitore di buone intenzioni e idee vagamente definite, piuttosto che in un utile strumento.

Va inoltre sottolineato che molte lingue europee non possiedono un'esatta corrispondenza al termine di origine anglosassone. In una prima accezione, largamente condivisa nella teoria e nella pratica, l'*accountability* può essere definita come il processo attraverso il quale un soggetto è chiamato a giustificare le proprie azioni o decisioni a un'autorità designata.

In questa prospettiva, l'*accountability* è esterna, poiché implica che la rendicontazione avvenga nei confronti di soggetti o enti esterni rispetto a chi è chiamato a risponderne. Inoltre, essa presuppone una relazione sociale e implica un diritto da parte di chi richiede spiegazioni, accompagnato dall'accettazione di eventuali critiche o sanzioni da parte di chi risponde. Un esempio immediato di tale dinamica può essere individuato nel rapporto tra cittadini e

rappresentanti politici, in cui gli elettori cercano di responsabilizzare i propri rappresentanti rispetto alle promesse elettorali, alle decisioni adottate e alle azioni intraprese.

È importante sottolineare che questa interpretazione tende a identificare il concetto di *accountability* con quello di responsabilità, benché quest'ultimo possa essere inteso in molteplici modi diversi[201]. Tale enfasi - storicamente tipica, sebbene con diverse *nuances*, del contesto anglosassone - ha condotto a una visione in cui l'*accountability* è essenzialmente legata al controllo esterno e, solo in secondo luogo, alla responsabilità interna di un individuo verso i propri valori morali o la propria coscienza[202]. Tuttavia, questa focalizzazione copre solo una parte delle attività e dei processi che implicano un certo grado di responsabilità. Di conseguenza, il concetto si è nel tempo ampliato, spingendo verso una riduzione dei confini

[201] MULGAN, R. (2002), "'Accountability': an ever-expanding concept?", *Public administration*, 78(3), pp. 555-573, https://doi.org/10.1111/1467-9299.00218.

[202] FINER, H. (2018), "Administrative responsibility in democratic government", *Classics of administrative ethics*, Routledge, pp. 5-26.

della responsabilità (*responsibility*) e un'estensione di quelli dell'*accountability*[203].

Il termine *accountability* è stato, difatti, adattato e ampliato per includere nuove situazioni e contesti. Un primo ampliamento lo associa a un senso di responsabilità personale e a una preoccupazione per l'interesse pubblico. Questa visione c.d. interna si allontana dall'enfasi esclusiva sul controllo esterno e si integra con l'etica professionale e personale. L'idea centrale è che maggiore fiducia nel giudizio e nella moralità degli operatori possa ridurre la necessità di controlli esterni, sempre più pressanti[204].

Un'altra interpretazione[205] assimila pienamente l'*accountability* ai sistemi di controllo democratici utilizzati per monitorare l'operato dei governi. In questa visione, garantire l'*accountability* significa creare un'architettura istituzionale che permetta un adeguato controllo sui soggetti pubblici.

[203] DUBNICK, M. (2012), "Clarifying accountability: An ethical theory framework" in *Public sector ethics*, Routledge, pp. 68-81.

[204] DUBNICK, M. (2012), op. cit.

[205] FRIEDRICH, C. J. (2000), "Public policy and the nature of administrative responsibility" in *The Science of Public Policy* (1ª edizione 1940).

Un'ulteriore espansione del concetto riguarda la *responsiveness*, ovvero la capacità dei governi di rispondere alle necessità e ai desideri dei cittadini, indipendentemente da pressioni coercitive. A differenza del controllo esterno, la *responsiveness* si basa su una relazione non coercitiva tra cittadini e istituzioni[206].

Infine, sempre in accordo con i citati studi, il termine *accountability* è stato usato per descrivere il dibattito pubblico tra cittadini e governi, che è alla base delle democrazie partecipative. Anche in assenza di autorità formali o di relazioni gerarchiche, *l'accountability* in questo contesto si concentra sul mantenimento di un dialogo partecipativo tra le parti coinvolte.

Le definizioni descritte finora trovano ampio consenso sia nella teoria che nella pratica. Tuttavia, alcune posizioni differiscono. Due contributi significativi a riguardo provengono dalla Commissione Europea e dallo standard *AccountAbility Principles* (AA1000)[207].

[206] MULGAN, R. (2002), art. cit.

[207] Si veda sul punto il documento *European Governance: A White Paper*, COM (2001) 428, 25 luglio 2001. Inoltre, l'AA1000 (2008, https://www.accountability.org/standards) è un *framework* che mira a chiarire come i principi di responsabilità e sostenibilità siano connessi e complementari. L'idea centrale è che l'*accountability* sia il principio alla base della rendicontazione finanziaria e di altre forme di *accounting*, *auditing* e *reporting*. Si enfatizza, tuttavia, che nel rapporto tra responsabilità e sostenibilità, permanga scarsa chiarezza.

La Commissione Europea - concentrandosi sulle implicazioni sociopolitiche delle relazioni tra Stati, istituzioni e cittadini - adotta una visione che si colloca tra la *responsibility* e la *responsiveness*. L'UE considera l'*accountability* uno dei principi fondamentali, sottolineando l'importanza di una trasparenza sempre maggiore nei processi decisionali e legislativi. Ogni istituzione è chiamata a spiegare e giustificare le proprie azioni, garantendo chiarezza e responsabilità a livello sia europeo che nazionale[208].

Lo standard AA1000, invece, è rivolto alle aziende e si concentra sulla sostenibilità. Secondo questo approccio, l'*accountability* include la consapevolezza e la responsabilità rispetto all'impatto delle politiche e delle decisioni aziendali. Le organizzazioni sono chiamate a coinvolgere gli stakeholder nella definizione e comprensione delle questioni rilevanti, a rispondere alle loro aspettative e a fornire informazioni trasparenti sulle proprie azioni e *performance*.

In sintesi, il concetto di *accountability*, con radici profonde nella filosofia morale, nel diritto e nella gestione organizzativa, pur

[208] PARLAMENTO EUROPEO, Le procedure decisionali di natura intergovernativa, Note tematiche sull'Unione europea, 2024, https://www.europarl.europa.eu/factsheets/it/sheet/9/responsabilita-e-trasparenza-nell-ue.

divenuto sempre più articolato, integrando dimensioni interne ed esterne, etiche e istituzionali, oltre che legate alla partecipazione democratica, può essere in sintesi compreso con riguardo ai suoi pilastri strutturali:

- Chiarezza della responsabilità: identificazione chiara di chi è responsabile per una determinata azione o decisione.
- Controllo e verifica: sistemi che monitorano e valutano performance e decisioni.
- Meccanismi di riparazione: procedure che garantiscono che errori o abusi siano corretti e che le parti danneggiate ricevano compensazione.

Tuttavia, con l'introduzione dell'intelligenza artificiale, questi principi devono essere rivisitati per affrontare le nuove complessità introdotte dai sistemi autonomi e algoritmici. L'IA porta una trasformazione in questi concetti chiave:

- Ambiguità delle responsabilità: nei sistemi di IA, la responsabilità è distribuita (condivisione della responsabilità) tra sviluppatori, fornitori di dati, organizzazioni e utenti finali, complicando l'attribuzione

diretta della responsabilità. Ciò crea un livello di interdipendenza mai visto prima.

- Opacità decisionale: l'uso di modelli di *deep learning* e algoritmi complessi rende difficile comprendere come e perché una decisione è stata presa, sfidando i tradizionali meccanismi di controllo e verifica. Ciò è noto come il problema della "*black box*" (il filone dell'*explainable* IA consente, attraverso diverse tecniche, di rendere maggiormente "trasparenti" gli algoritmi di IA)
- Nuove forme di riparazione: gli errori generati dall'IA richiedono approcci innovativi per identificare e risolvere le conseguenze negative, come bias algoritmici o violazioni della *privacy*. Eventuali forme per ovviare alle conseguenze negative includono revisioni retroattive e meccanismi di risarcimento pre-programmati.

La natura distribuita dell'IA modifica anche il concetto di relazione tra gli attori, introducendo nuovi livelli di interdipendenza. Ad esempio, il ruolo dei fornitori di dati diventa centrale, poiché la qualità dei dati influenza direttamente il funzionamento dell'IA. Inoltre, emerge la figura del "titolare della trasparenza",

responsabile di comunicare le decisioni e le limitazioni del sistema a stakeholder e utenti finali. Questi cambiamenti pongono una pressione significativa sulle organizzazioni per sviluppare nuove competenze e strumenti per gestire la responsabilità in un ecosistema tecnologico complesso.

Alla luce di quanto precede, si può sostenere in linea generale che l'*accountability* nell'era dell'IA, dunque, richieda l'introduzione di principi e strumenti specifici:

- Responsabilità condivisa: l'*accountability* deve essere distribuita lungo l'intera catena di sviluppo e utilizzo dell'IA, con ogni attore che assume responsabilità proporzionali al proprio ruolo. Ciò richiede una mappa chiara delle responsabilità, accompagnata da contratti e accordi che specificano i ruoli. Inoltre, vanno previsti meccanismi semplificati sia per individuare l'attore responsabile, sia per dimostrare il nesso casuale tra danno e Sistema IA, nel caso in cui la parte lesa sia il consumatore finale, solitamente parte debole del sinallagma contrattuale.
- Documentazione e tracciabilità: ogni fase del ciclo di vita dell'IA, dalla progettazione alla distribuzione, deve essere

documentata per garantire la tracciabilità delle decisioni. Sistemi di *audit* e *log* digitali possono essere utilizzati per registrare ogni decisione critica.

- Supervisione umana: nonostante l'autonomia dell'IA, è essenziale che gli esseri umani mantengano la supervisione e possano intervenire in caso di malfunzionamenti. Ciò richiede l'adozione di meccanismi di "*kill switch*" per interrompere operazioni potenzialmente dannose.
- Standardizzazione: le aziende devono adottare standard comuni per garantire che la progettazione e la messa in opera dell'IA rispettino criteri etici e tecnici. Organizzazioni internazionali come ISO e IEEE stanno lavorando su standard globali per l'uso responsabile dell'IA.
- Formazione e sensibilizzazione: le organizzazioni devono investire in formazione per sviluppare una cultura aziendale orientata alla responsabilità etica e tecnica nell'uso dell'IA. Ciò include non solo il personale tecnico, ma anche i dirigenti e gli utenti finali.

Ulteriore **elemento che esplica esigenze di contemperamento**, con riferimento all'*accountability* nell'era dell'IA, **è quello della**

trasparenza. **La trasparenza, dimensione strutturale e strumentale agli obiettivi di *accountability***, implica tradizionalmente la capacità di comprendere e verificare le decisioni e i processi e nell'IA, la trasparenza deve fronteggiare sfide uniche:

- Spiegabilità (*explainability*): le decisioni basate su IA devono essere spiegabili, anche quando utilizzano modelli complessi come le reti neurali. Ciò richiede strumenti che traducano le decisioni algoritmiche in termini comprensibili. **La spiegabilità è cruciale per settori regolamentati come finanza e sanità**, dove le decisioni devono essere giustificabili agli occhi di regolatori e clienti o anche negli ambiti legati alla gestione delle risorse umane e alla giustizia.
- Accessibilità: la trasparenza richiede che le informazioni siano comprensibili non solo agli esperti, ma anche agli utenti finali e alle parti interessate. Ad esempio, un cliente bancario deve comprendere come un algoritmo abbia valutato il suo merito creditizio, senza doversi confrontare con un linguaggio tecnico incomprensibile.
- Comunicazione chiara: le aziende devono dichiarare apertamente i limiti dei sistemi IA, come i bias nei dati di addestramento o le incertezze nelle predizioni. La

consapevolezza di tali limiti è essenziale per prevenire aspettative irrealistiche e promuovere un uso responsabile.

La tensione tra trasparenza e confidenzialità, poi, rappresenta un'ulteriore criticità. Mentre le organizzazioni devono proteggere i segreti commerciali e i dati sensibili, è fondamentale fornire informazioni sufficienti per garantire responsabilità e fiducia.

La soluzione potrebbe risiedere nella c.d. trasparenza differenziata[209], per la quale informazioni dettagliate sono condivise con autorità di regolamentazione e stakeholder selezionati, mentre una spiegazione semplificata è fornita al pubblico generale. Strumenti come *report* di *audit* pubblici e dichiarazioni di responsabilità possono migliorare la trasparenza senza compromettere la confidenzialità.

Nel contesto dei *social media*, ad esempio, la trasparenza è essenziale per spiegare come vengono personalizzati i feed degli utenti, riducendo il rischio di manipolazioni psicologiche o discriminazioni algoritmiche. In aggiunta, la spiegabilità dei

[209] L. MACVITTIE, "I due concetti cruciali dell'Intelligenza Artificiale: trasparenza ed *explainability*", *Network 360*, 25 ottobre 2024, https://www.esg360.it/smart-technology/i-due-concetti-cruciali-dellintelligenza-artificiale-trasparenza-ed-explainability/

sistemi di IA sta avanzando grazie a metodi tecnici innovativi, come le tecniche di interpretabilità basate su *feature importance*.

Questi strumenti consentono di identificare quali variabili hanno avuto il maggiore impatto su una decisione algoritmica, migliorando la comprensione e la fiducia nei sistemi IA complessi. Ad esempio, tecniche come LIME (*Local Interpretable Model-Agnostic Explanations*) e SHAP (*SHapley Additive exPlanations*) sono state sviluppate per offrire spiegazioni dettagliate e comprensibili, anche in ambiti come il credito al consumo e la diagnostica medica.

Il complesso quadro concettuale sopra sintetizzato, peraltro, deve essere interpretato anche alla luce di una serie di questioni etiche che vanno oltre le sfide tradizionali dell'*accountability* e della trasparenza e che rappresentano sfide e rischi emergenti per i contesti di *business*.

Come discusso, l'IA può porre temi di bias e discriminazione. Gli algoritmi possono perpetuare o amplificare pregiudizi esistenti, con implicazioni sociali significative. Ad esempio, in

linea teorica, sistemi di selezione del personale basati su IA potrebbero favorire inconsapevolmente determinati gruppi demografici. Mitigare questo rischio richiede una combinazione di monitoraggio continuo, diversificazione dei dataset e interventi correttivi nei modelli.

Il tema della *privacy* e della sorveglianza, poi, pone un certo grado di preoccupazione poiché **l'utilizzo di dati personali per addestrare i modelli di IA apre questioni di non poco conto in relazione alla protezione e anonimizzazione dei dati e al consenso informato**. Altresì, esistono alcuni interrogativi tuttora irrisolti con riguardo al pericolo di manipolazione.

L'IA può essere utilizzata per influenzare il comportamento umano, sollevando interrogativi sull'autonomia individuale e sull'integrità delle decisioni. Difatti, gli algoritmi di raccomandazione potrebbero favorire determinate scelte senza che gli utenti ne siano del tutto consapevoli, o meglio, gli utenti potrebbero non comprendere fino in fondo il motivo della loro decisione, che da un certo punto di vista potrebbe non sembrare del tutto libera. Non di

meno, pure il tema dell'equità è in discussione. Le decisioni automatizzate devono (dovrebbero) rispettare i principi di equità e giustizia, evitando disparità di trattamento tra gruppi diversi. Ciò implica una verifica costante dei criteri utilizzati nei modelli decisionali e la revisione delle metriche di *performance* per garantire l'equilibrio tra accuratezza ed equità.

Quanto sopra - in termini di ripercussione sulla sfera aziendale - determina un ampio novero di rischi, nuovi o rivisitati. In estrema sintesi, difatti, sono oggetto di attenzione gli aspetti legati alla *compliance*: l'adozione dell'IA richiede alle aziende di rispettare normative complesse, come il GDPR e lo stesso AI Act in Europa, che impongono standard rigorosi per la trasparenza e la protezione dei dati.

Le violazioni possono comportare sanzioni significative, come nel caso di multe inflitte per l'uso improprio di dati personali. Crescenti sono poi le preoccupazioni legate al rischio (danno) reputazionale derivante da errori o le violazioni etiche legati all'uso dell'IA che possono danneggiare gravemente la

reputazione aziendale, portando a una perdita di fiducia da parte di clienti, investitori e dipendenti.

Le aziende, oltre all'adeguamento a standard rigorosi, dovranno affrontare crescenti investimenti anche in comunicazione trasparente per mitigare questi rischi, altresì andando incontro a possibili costi in caso di contenziosi derivanti da decisioni errate o discriminatorie prese dall'IA che possono comportare ripercussioni significative e sanzioni. Occorre tenere conto, poi, che l'esigenza di rispettare standard elevati di *accountability* e trasparenza potrebbe aumentare i costi e i tempi di sviluppo, potendo tutto ciò in linea teorica rallentare l'introduzione di nuovi prodotti e servizi imperniati su IA con possibili vantaggi competitivi a lungo termine per quegli operatori che risponderanno in modo "*proactive*" alle sfide che essa pone attraverso soluzioni con caratteristiche rispondenti a una IA responsabile e sostenibile.

Tali sintetiche e generali considerazioni possono essere meglio comprese soffermandosi su alcuni settori, cosiddetti 'sensibili'.

In linea genale si ricorda che per il settore sanitario gli errori nei sistemi di IA per la diagnosi potrebbero avere conseguenze gravi per i pazienti. I modelli di IA devono essere addestrati su dati rappresentativi per evitare diagnosi inefficaci. Ciò implica stretta collaborazione tra istituzioni sanitarie, personale medico e sviluppatori di IA per migliorare la qualità dei dataset di addestramento e dare vita a sistemi di IA che possano essere di sostegno all'operatore sanitario (senza sostituirlo e nel rispetto del principio "*human in the loop*").

In tal senso, investire in formazione sull'IA può migliorarne l'adozione responsabile nel contesto medico. Tuttavia - sebbene la trasparenza sia fondamentale per costruire fiducia e garantire l'accettazione da parte di medici e pazienti - l'uso di dati sensibili richiede un bilanciamento tra trasparenza e protezione dei dati, con implicazioni legali significative. Le aziende devono quindi adottare tecniche come l'anonimizzazione per proteggere i dati al contempo garantendo l'efficacia della informazione condivisa.

Una esemplificazione teorica interessante potrebbe poi essere ricondotta al settore della giustizia. I sistemi di IA eventualmente

utilizzati per determinare cauzioni, sentenze o probabilità di recidiva devono essere altamente trasparenti per evitare decisioni arbitrarie o discriminatorie. I dati storici utilizzati per addestrare i modelli possono contenere pregiudizi sistemici, perpetuando ingiustizie. È essenziale che le decisioni prese dall'IA siano soggette a revisione umana per garantire equità e conformità ai principi legali. Ciò può includere comitati di revisione multidisciplinari. Inoltre, le autorità devono introdurre regolamenti che richiedano *audit* obbligatori e le organizzazioni devono investire in strumenti di mitigazione dei bias.

Ancora, con riferimento alla finanza, sempre a titolo di esempio teorico, la mancanza di trasparenza nei modelli di IA utilizzati per valutare l'affidabilità creditizia potrebbe portare a discriminazioni e violazioni normative. Gli errori nei modelli predittivi potrebbero influenzare interi mercati finanziari, con ripercussioni significative per l'economia globale. Adottare strumenti di spiegabilità può migliorare la fiducia degli utenti. Inoltre, sempre riferito all'IA, è prevista l'introduzione di standard di gestione del rischio e la richiesta di *audit* regolari dei suoi modelli per garantire *compliance*

e prevenire abusi. Le aziende devono collaborare con i regolatori per sviluppare linee guida comuni.

In conclusione, l'*accountability* e la trasparenza nell'IA rappresentano sfide complesse ma fondamentali per garantire un uso etico e responsabile di queste tecnologie. Le aziende devono adottare un approccio proattivo, investendo in strumenti e pratiche che bilancino le esigenze di innovazione con i principi etici e normativi. In settori ad alta complessità come giustizia, sanità e finanza, la capacità di garantire responsabilità e spiegabilità può determinare il successo o il fallimento di un'iniziativa fondata su di essa, con implicazioni profonde per la società nel suo complesso.

IV.2. *EXPLAINABILITY* E STRUMENTI PER UN CONTROLLO ALGORITMICO EFFICACE

Dato il sempre più frequente uso di strumenti di IA, sia nella sfera privata sia in quella lavorativa, diventa fondamentale prendere in considerazione un aspetto di questa nuova tecnologia, cioè la spiegabilità.

È risaputo come l'IA agisca attraverso l'uso di algoritmi che, nel migliore dei casi, sono chiari solamente a chi effettivamente li scrive; per la maggior parte delle persone è impensabile capire e comprendere i meccanismi, i collegamenti, le inferenze statistiche che hanno portato un dato risultato. Mentre, in prima analisi, si potrebbe sorvolare su questo aspetto - considerando che ognuno di noi può analizzare il risultato senza per forza subirlo passivamente e senza dover comprendere a pieno il percorso seguito - va considerato il fatto che capire il processo che ha portato da un *input* ad un *output*, può permetterci di assicurarci che l'IA persegua altri principi fondamentali, già presenti nei principi fondamentali della bioetica, quali beneficenza, non maleficenza, giustizia e autonomia.

La possibilità, dunque, di comprendere cosa c'è dietro quella che comunemente viene definita "*black-box*" diventa un valore, non solo aggiunto ma fondamentale, così come, a cascata, la possibilità di intervenire a correzione degli algoritmi, laddove non rispettino più i principi fondamentali menzionati poco fa.

Inoltre, comprendere le logiche e le motivazioni di un *output* proveniente da un algoritmo è fondamentale per garantire il

rispetto della dignità umana, in modo che non possa essere possibile subire una decisione senza averne chiaro il *quid*.

Per affrontare tale sfida, è sempre più importante parlare di intelligenza artificiale spiegabile (*explainable* AI o XAI), che si propone di rendere i modelli di IA più trasparenti e comprensibili, facilitandone il monitoraggio. Questo approccio permette agli utenti di comprendere le logiche sottostanti alle decisioni algoritmiche, trasformando ciò che oggi è percepito come un meccanismo imperscrutabile in un sistema più accessibile e verificabile. Questa trasparenza non solo è funzionale ad aumentare la *confidence* nell'approcciare questi strumenti, ma diventa fondamentale per individuare potenziali bias o errori.

Un passo decisivo in questa direzione è rappresentato dall'introduzione di tecniche di monitoraggio continuo degli algoritmi, che consentano di supervisionare costantemente le prestazioni e gli output generati. Rilevare tempestivamente eventuali deviazioni o anomalie nel comportamento degli algoritmi permette non solo di migliorarne l'efficacia, ma anche di prevenire l'insorgenza di problematiche etiche o legali.

L'integrazione di strumenti di monitoraggio con pratiche di *audit* algoritmico inoltre rafforza la capacità di controllo delle organizzazioni, assicurando che l'IA non si limiti a produrre risultati, ma lo faccia in modo conforme ai principi di trasparenza e responsabilità, facilitando l'adozione di questi strumenti.

Un ulteriore aspetto da considerare nel garantire l'affidabilità e l'equità dei modelli di intelligenza artificiale riguarda la gestione dei dati utilizzati durante la fase di addestramento. Dati distorti o non più rappresentativi possono introdurre bias nei modelli, portando a decisioni discriminatorie o inesatte. Per affrontare questa sfida, è fondamentale adottare processi rigorosi di pulizia e selezione dei dati, assicurando che il modello sia costruito su una base solida e priva di distorsioni.

Anche con le migliori pratiche di gestione dei dati, possono emergere situazioni in cui è necessario rimuovere specifici dati dal modello già addestrato, ad esempio perché attraverso il monitoraggio dell'algoritmo è emerso come un determinato dataset potrebbe contenere informazioni fuorvianti o discriminatorie.

In questi casi, il "*machine unlearning*" si propone come una soluzione efficace[210]. Questa disciplina, ancora in fase di studio e soggetta alle limitazioni in termini di affidabilità del caso, sviluppa metodi che consentono di eliminare in modo efficiente ed efficace determinati dati da un modello addestrato, senza la necessità di un riaddestramento completo, riducendo così l'impatto in termini di tempi e costi e rendendo possibile, ad esempio, rimuovere dati personali dai modelli al fine di tutelare la *privacy* degli individui e di migliorare la sicurezza dei modelli.

IV.3. DISCRIMINAZIONE ALGORITMICA E BIAS NEI MODELLI DI IA

Un aspetto da sempre centrale nell'uso dell'intelligenza artificiale riguarda il rischio di bias nei modelli algoritmici. Con il termine bias si fa riferimento a distorsioni negli output, spesso provocate da pregiudizi umani inconsapevolmente infusi nei dati di addestramento o nei criteri con cui l'IA elabora le decisioni, quindi nell'algoritmo stesso. Poiché l'IA è sempre più integrata in processi decisionali determinanti, questo problema

[210] SAI, S., MITTAL, U., CHAMOLA, V. *ET AL*, "Machine Un-learning: An Overview of Techniques, Applications, and Future Directions", *Cogn Comput* 16, 2024, pp. 482–506, https://doi.org/10.1007/s12559-023-10219-3.

non può essere trascurato e richiede un'analisi approfondita delle cause e dei possibili metodi di mitigazione.

I modelli di IA sono costruiti su principi logici e matematici predefiniti, attraverso cui elaborano e apprendono dai dati forniti. Questa necessità di basarsi su pattern logici predeterminati e costanti, rischia di limitare la capacità dell'algoritmo di comprendere a pieno il contesto in cui opera, soprattutto quando è applicato a scenari differenti da quelli per cui è stato progettato. A ciò si aggiunge il fatto che la principale fonte di bias nei modelli, risiede nei dati con cui questi sono allenati. La raccolta e il campionamento dei dati sono, per loro natura, processi soggetti a scelte umane e quindi, a bias cognitivi più o meno marcati, e quasi sempre non sottoposti ad un rigoroso metodo scientifico. In molti casi, le discriminazioni presenti nel dataset non sono nemmeno immediatamente riconoscibili, rendendo ancora più complesso il problema.

Come detto poc'anzi, se i dati utilizzati per addestrare un modello presentano discriminazioni sociali o stereotipi, l'algoritmo sarà portato a considerarli e utilizzarli nella produzione dei risultati o nell'assunzione di decisioni,

contribuendo a rafforzare esempi di esclusione o disparità già presenti nella società.

Questo fenomeno può avere conseguenze importanti, soprattutto su gruppi vulnerabili. Un esempio emblematico è stato osservato negli algoritmi di selezione del personale, che hanno mostrato tendenze discriminatorie nel valutare le candidature sulla base di dati storici sbilanciati a favore di assunzioni maschili. Allo stesso modo, alcuni sistemi di valutazione del rischio impiegati in ambito giudiziario e sociale hanno dimostrato alcuni limiti, che potrebbero influenzare negativamente le decisioni riguardanti la protezione di alcune categorie particolarmente esposte.

Sono situazioni che evidenziano come questo tipo di bias non solo possano ricalcare, ma addirittura amplificare, le disuguaglianze esistenti. Per tal motivo, è essenziale adottare strategie per ridurre il bias nei modelli di IA. **La diversificazione dei dataset di addestramento è un primo passo fondamentale: includere dati rappresentativi di diverse popolazioni e categorie contribuisce a ridurre le distorsioni**.

Tecniche di pre-elaborazione possono inoltre identificare e correggere eventuali squilibri prima che i dati vengano utilizzati dal modello. Però, tutto ciò non basta: come detto in precedenza è fondamentale un controllo costante dell'IA, per verificare che le decisioni prodotte non mostrino anomalie, ingiustizie o incoerenze rispetto agli obiettivi per cui è stata adottata.

Torna dunque l'importanza di integrare sempre di più tecniche che rendano gli algoritmi spiegabili e quindi più semplici da monitorare e correggere, ad esempio grazie al già citato approccio del "*machine unlearning*". Definire al più presto standard condivisi - per esempio tramite future certificazioni ISO specifiche per l'intelligenza artificiale - consentirebbe di garantire presìdi realmente conformi, tagliare i costi di doverli progettare da zero, uniformare le tutele in tutta l'Unione europea con benefici anche competitivi e sostenere in particolare le piccole e medie imprese, che altrimenti faticherebbero a sostenere un processo tanto impegnativo.

Anche la composizione dei gruppi di sviluppo degli algoritmi gioca un ruolo fondamentale. Questi non sono neutrali: rispecchiano le scelte e le prospettive di chi li progetta. Integrare competenze diverse nei processi di sviluppo dell'IA, che vadano dalle scienze computazionali alla sociologia, fino all'etica e alla giurisprudenza, aiuta a individuare e correggere pregiudizi che altrimenti potrebbero passare inosservati. Promuovere una cultura della responsabilità all'interno delle organizzazioni è altrettanto importante: l'adozione di protocolli di valutazione del rischio, *audit* algoritmici e processi di supervisione trasparente può fare la differenza tra un sistema equo e uno potenzialmente dannoso.

Solo attraverso un approccio strutturato, che combini dataset diversificati, monitoraggio continuo e processi di correzione mirati, è possibile sviluppare sistemi di IA più equi, inclusivi e affidabili. L'IA deve essere uno strumento che abilita opportunità, senza rafforzare le discriminazioni esistenti. Garantire che operi in modo giusto e responsabile non è solo una scelta tecnica, ma un'esigenza etica e sociale.

IV.4. *PRIVACY*, SICUREZZA DEI DATI

Insomma, come già evidenziato, lo sviluppo e l'adozione di sistemi di intelligenza artificiale accelera il cambiamento nei settori più disparati, con impatti profondi e trasformativi. Così non solo si affermano nuove opportunità di *business*, modelli operativi più efficienti, nuovi consumi e nuove esperienze ad essi legate, ma anche nuovi modi di intendere libertà, diritti e doveri, nei quali l'idea stessa di cittadinanza tradizionalmente intesa trova una nuova e ulteriore dimensione in quella digitale.

Se, in via generale, è possibile sostenere che le aziende e i governi che hanno una spiccata attitudine al cambiamento tecnologico hanno un sensibile vantaggio, tanto più nell'adottare soluzioni IA avranno un vantaggio competitivo cruciale in un mondo sempre più digitalizzato e interconnesso nel quale la protezione dei dati e la sicurezza informatica diventano sfide globali, cruciali e non eludibili.

La raccolta di dati personali per alimentare i modelli IA offre opportunità straordinarie, tra le quali: la personalizzazione dei servizi, il miglioramento delle *performance* dei modelli stessi,

l'innovazione nei settori più diversi, il sostegno decisionale in ambito pubblico e privato, la semplificazione burocratica, lo sviluppo di nuovi modelli di *business*.

Tuttavia, può comportare anche rischi significativi per la *privacy* e la sicurezza, con risvolti sulla fiducia dei consumatori e la competitività aziendale. In aggiunta a ciò, si consideri che la condivisione dei dati (*data sharing*) rappresenta un elemento cruciale per affrontare sfide transnazionali come il riciclaggio di denaro, il finanziamento del terrorismo e la criminalità organizzata. Eppure, questa esigenza di cooperazione richiede di essere contemperata con i principi di protezione dei dati personali e i vincoli normativi sulla *privacy*.

La tensione tra sicurezza e tutela dei diritti individuali è particolarmente evidente nei processi di scambio informativo tra autorità di paesi diversi, i quali spesso operano secondo quadri giuridici eterogenei e, talvolta, inconciliabili. Se, da una parte, strumenti come la Direttiva antiriciclaggio dell'Unione Europea e i regolamenti del *Financial Action Task Force* (FATF) promuovono una maggiore trasparenza e tracciabilità delle

transazioni finanziarie; dall'altra, il Regolamento generale sulla protezione dei dati (GDPR) impone severe limitazioni alla diffusione transfrontaliera di dati personali.

Ne deriva un potenziale rallentamento delle indagini internazionali e una compromissione dell'efficacia degli strumenti di *intelligence* economico-finanziaria.

L'assenza di un sistema armonizzato e universalmente accettato per la condivisione dei dati, in grado di bilanciare efficacemente la necessità di proteggere i diritti fondamentali con quella di garantire la sicurezza pubblica. rappresenta una delle maggiori criticità.

Diventa allora imprescindibile sviluppare meccanismi interoperabili, fondati su standard condivisi di sicurezza, trasparenza, proporzionalità e *accountability*, capaci di favorire una cooperazione leale tra Stati, autorità pubbliche e soggetti privati, nel rispetto delle diverse sovranità normative.

Il dibattito sul *data sharing* evidenzia così la necessità di nuovi paradigmi giuridici e tecnologici, per integrare in modo

equilibrato le finalità di contrasto al crimine con un approccio etico e multilivello alla *governance* dei dati, tenendo conto delle asimmetrie geopolitiche e delle diverse sensibilità culturali in materia di *privacy*.

Si vuole qui esaminare come le vulnerabilità nei sistemi di IA possano essere sfruttate in attacchi cibernetici, esplorando la crescente complessità della *governance* globale dei dati. Si analizza quindi il panorama normativo frammentato, le sfide in materia di sovranità dei dati e le implicazioni per le politiche *antitrust* e come questi fattori possano influenzare le strategie aziendali. In un mondo in cui la protezione dei dati, *l'accountability* e la trasparenza sono diventate priorità assolute per garantire un uso etico e responsabile di queste tecnologie, si vuole qui illustrare quali siano i principali elementi a caratterizzare questo complesso quadro. La gestione responsabile dei dati e la sicurezza informatica sono, infatti, diventate non solo necessità operative, ma anche determinanti strategici per il successo globale.

L'intelligenza artificiale sta rapidamente trasformando settori economici, sociali e politici, con implicazioni straordinarie per le modalità in cui i dati sono raccolti, trattati e protetti. I sistemi di IA si fondano su enormi volumi di dati, molti dei quali di natura personale, che sono utilizzati per addestrare modelli di *machine learning* e altre tecnologie predittive. Mentre l'uso di dati personali consente di migliorare l'efficacia e la precisione di tali sistemi, solleva anche preoccupazioni profonde in materia di *privacy*, sicurezza e *governance* globale.

In questo contesto, la protezione dei dati e la sicurezza cibernetica si intrecciano con la responsabilità delle imprese, la regolamentazione *antitrust* e la gestione multilivello dei dati in un quadro geopolitico sempre più frammentato.

In via preliminare è opportuno considerare che la raccolta di dati personali è il cuore dello sviluppo dei sistemi di IA. Tali dati possono includere informazioni personali identificabili, comportamentali, sanitarie, finanziarie e altro ancora, raccolti da dispositivi connessi, applicazioni *web*, *social media* e piattaforme di *e-commerce*.

Il processo di raccolta dei dati - seppur cruciale per l'innovazione - solleva delicate questioni riferibili a controlli indiscriminati di massa, alla profilazione dettagliata e allo *scoring*, alla manipolazione comportamentale e alla vulnerabilità di persone o di gruppi di persone, ed espone le aziende e le autorità governative al rischio di abusi a detrimento di diritti riconosciuti come fondamentali dalle costituzioni liberali, dalla dichiarazione universale dei diritti umani e, da questa parte di mondo, dalla Carta dei diritti fondamentali dell'Unione europea.

Si rinvia al successivo capitolo V lo sguardo e l'approfondimento del quadro normativo globale e frammentato dal quale emergono delicati e sfaccettati aspetti di *governance*, qui è invece opportuno segnalare la preoccupazione per la sicurezza informatica e dei dati in quanto i potenziali rischi correlati sono esacerbati dalla sofisticazione degli attacchi cibernetici e dalla vulnerabilità di molte delle tecnologie emergenti ed esponenziali.

Le minacce alla sicurezza informatica e dei dati assumono una rilevanza ancora maggiore all'interno di un contesto internazionale globale e interconnesso. I cyberattacchi non sono solo l'opera di gruppi di criminali informatici, ma anche di attori, direttamente o indirettamente riferibili ad apparati governativi che utilizzano la tecnologia per perseguire obiettivi strategici.

Le statistiche globali sugli attacchi informatici mostrano una crescita significativa sia di quelli perpetrati da criminali informatici sia di quelli attribuibili, direttamente o indirettamente, a governi, sebbene la distinzione tra attacchi perpetrati da criminali informatici e quelli "commissionati" da governi non è sempre netta, chiara, e dimostrabile.

Gli attacchi informatici rivolti a enti governativi sono aumentati più del 90% a livello globale. Nel primo semestre del 2024, sono stati registrati 1.637 attacchi informatici gravi a livello mondiale, con una media di 273 attacchi al mese, segnando un incremento del 23% rispetto allo stesso periodo dell'anno precedente. La sanità è risultata il settore più colpito a livello globale. Si stima che nel

2024, il costo globale del crimine informatico abbia raggiunto i 10.000 miliardi di euro, il doppio rispetto all'anno precedente[211].

Le tecnologie di IA, in particolare i sistemi di raccolta e analisi dei dati, sono al centro di queste operazioni. Gli attacchi informatici, sia nei confronti di entità governative che di aziende private, possono compromettere i modelli di IA e sfruttare vulnerabilità critiche nei sistemi di sicurezza dei dati.

In termini di *governance*, da una parte, l'*accountability* e la trasparenza, e dall'altra la tutela della concorrenza e della pluralità di mercato, oltre che la tutela di diritti fondamentali, richiedono un approccio multilivello in grado di coinvolgere attori nazionali e internazionali. La *governance* globale dei dati è uno dei terreni di maggiore confronto per la sovranità tecnologica e rappresenta una questione sempre più centrale nel panorama geopolitico attuale.

[211] ASSOCIAZIONE ITALIANA PER LA SICUREZZA INFORMATICA, *Rapporto Clusit 2024*, ottobre 2024, https://clusit.it/pubblicazioni/; ISTITUTO PER LA COMPETITIVITÀ, "Il settore sanitario è il principale bersaglio degli attacchi informatici", 24 gennaio 2025, https://www.i-com.it/2025/01/24/il-settore-sanitario-e-il-principale-bersaglio-degli-attacchi-informatici-dalla-commissione-ue-un-piano-dazione-per-la-cybersecurity/.

La localizzazione dei dati - che implica l'obbligo di conservare i dati all'interno dei confini di uno specifico Stato - è una misura che alcuni paesi adottano per proteggere la *privacy* dei propri cittadini e limitare l'influenza da parte di altri. Tuttavia, tali misure possono ostacolare il libero flusso di dati e possono introdurre conflitti tra la necessità di proteggere i dati personali e quella di facilitare l'innovazione, lo sviluppo tecnologico, la crescita e lo sviluppo. Le organizzazioni sovranazionali stanno lavorando per sviluppare principi di *governance* condivisi, ma la diversità di interessi tra paesi sviluppati ed emergenti e tra attori privati e pubblici rende difficile raggiungere un consenso globale.

A livello geopolitico, la protezione dei dati assume in via crescente il rilievo di una questione di sicurezza nazionale, con impatti significativi per le relazioni internazionali e per la sovranità dei paesi. L'equilibrio tra innovazione tecnologica, *privacy*, sicurezza dei dati e concorrenza, richiede un impegno costante da parte dei governi, delle aziende e delle organizzazioni internazionali. La sfida della *governance* dei dati nell'era dell'IA apre quindi a considerazioni di ordine

costituzionale, transnazionale e globale, di politica economica e di giustizia sociale. Fenomeni tecnologici per dimensione e natura globale pongono una seria riflessione sull'impossibilità di essere regolato da norme o da ordinamenti per loro stessa dimensione e natura locali, al più continentali e comunque, non globali.

La vivacità del dibattito esteso al livello sovranazionale e animato da esperti, enti di ricerca e gruppi di lavoro istituiti in seno a governi nazionali, organismi e sedi internazionali (basti pensare al AI Act approvato in seno all'Unione europea, alla Dichiarazione di Parigi sulla Etica dell'IA (UNESCO, 2021), al *Global Digital Compact* e l'*AI Advisory Body* promossi dall'ONU, i Principi OCSE sull'IA del 2019 adottati anche dal G20, l'*AI Bill of Rights* degli Stati Uniti d'America, voluto dall'amministrazione Biden e revocato da quella Trump, l'"*Hiroshima Process*" lanciato dal G7 nel 2023 per regolare i modelli IA generativa, il *Global Partnership on AI* (GPAI) e le altre task force internazionali che promuovono linee guida e standard condivisi) e l'importanza più che simbolica attribuita alle iniziative legislative in materia di intelligenza artificiale, *privacy* e protezione dei dati, sicurezza

informatica e dei dati, si riflettono all'interno di un quadro regolatorio globale, eterogeneo e frammentato.

A seconda degli ordinamenti di riferimento i regolatori cercano di bilanciare necessità di investimento con esigenze di sicurezza e introducono nuovi standard che pongono questioni apprezzabili in termini economici e sostengono ambizioni di sovranità tecnologica, soprattutto con riferimento alle tecnologie emergenti ed esponenziali.

IV.5. UTILIZZO DELL'IA E IMPLICAZIONI SU *ANTITRUST*

Il **G7** delle autorità di concorrenza, riunito nel **comunicato del 4 ottobre 2024**[212], ha evidenziato le sfide emergenti nella gestione della concorrenza nell'era dell'Intelligenza Artificiale, in particolare per quanto riguarda le tecnologie di **Intelligenza Artificiale Generativa**. Il documento ribadisce l'importanza di garantire **mercati aperti, equi e contestabili** (*i.e.* contendibilità

[212] EUROPEAN COMMISSION, G7 Competition Summit: Effective international cooperation contributing to fair, open and contestable AI services, 4 October 2024, https://digital-markets-act.ec.europa.eu/g7-competition-summit-effective-international-cooperation-contributing-fair-open-and-contestable-ai-2024-10-04_en.

del mercato) per le tecnologie e i servizi legati all'IA, così da favorire l'ingresso di nuovi attori e stimolare l'innovazione.

L'IA, infatti, ha il potenziale di trasformare in modo profondo e positivo vari settori economici, aumentando la produttività e aprendo nuove opportunità di crescita. Tuttavia, il rischio di concentrazioni di potere di mercato è altrettanto rilevante. Il G7 ha identificato in particolare alcune caratteristiche dei mercati dell'IA che potrebbero favorire la centralizzazione del potere, come l'accesso limitato a risorse critiche (dati, chip, infrastrutture *cloud*, talenti) e l'adozione di pratiche anti-concorrenziali da parte delle grandi piattaforme tecnologiche, come il *self-preferencing*[213], il *tying*[214] e il *bundling*[215]. Questi fattori potrebbero ridurre la concorrenza, limitare le scelte dei consumatori e ostacolare l'ingresso di nuovi competitor nel mercato.

[213] *Self-preferencing*: condotta di una piattaforma dominante che conferisce sistematicamente un vantaggio competitivo ai propri prodotti o servizi rispetto a quelli dei terzi che vi operano.

[214] *Tying*: pratica con cui la vendita di un prodotto principale è vincolata all'acquisto obbligato di un secondo prodotto, limitando la libertà di scelta del cliente e potenzialmente escludendo la concorrenza sul mercato del prodotto vincolato.

[215] *Bundling*: Offerta congiunta di due o più prodotti in un unico pacchetto (solo in forma aggregata o anche separatamente), utilizzata per trasferire potere di mercato o ostacolare i concorrenti nei mercati collegati.

Un tema centrale del comunicato è l'impegno a promuovere un *enforcement antitrust* forte e tempestivo per garantire che i benefici derivanti dall'IA possano essere pienamente realizzati, a vantaggio di tutta l'economia e della società. In particolare, il G7 ha sottolineato l'importanza di politiche adattabili e lungimiranti, in grado di rispondere alle sfide specifiche derivanti dallo sviluppo delle tecnologie IA e di evitare che la concentrazione di potere minacci l'equità dei mercati.

Un'ulteriore riflessione in merito è stata espressa dal Direttore Generale della DG Concorrenza della Commissione Europea, Olivier Guersent, che ha di recente commentato la competitività del mercato dell'IA, affermando che attualmente vi sono più di trecento attori che operano in questo settore, ma che è importante vigilare su alcuni fattori chiave[216]. Sebbene il mercato sia competitivo, alcuni fattori cruciali, come i **chip specializzati** e le **infrastrutture *cloud***, sono sotto il controllo di grandi aziende tecnologiche. Fatto che potrebbe potenzialmente creare dei **colli**

[216] "The Future of Antitrust Enforcement & Policy", Concurrences events, Washington, April 1, 2025, https://www.concurrences.com/en/evenement/the-future-of-antitrust-enforcement-policy.

di bottiglia, che possono causare difficoltà all'ingresso di nuovi operatori e, quindi, per la concorrenza nel lungo periodo.

Guersent ha precisato che la Commissione Europea sta monitorando attentamente queste dinamiche per evitare il consolidamento del potere da parte di pochi soggetti, che potrebbero estendere il loro dominio dall'ambito digitale tradizionale a quello dell'IA, restringendo ulteriormente le opportunità per le piccole e medie imprese.

Un esempio concreto, e già discusso[217], di come l'IA possa influire sulla concorrenza è il caso del **cambiamento nel modello di *business* di Google con il suo *AI Overview***. Tradizionalmente, Google nasce come un motore di ricerca che indirizzava gli utenti verso siti esterni per trovare informazioni. Con l'introduzione di risposte generate dall'IA direttamente nella pagina dei risultati, Google ha modificato il suo approccio, rispondendo direttamente alle domande degli utenti, senza sempre un chiaro richiamo ai siti di notizie o contenuti originali. È un cambiamento significativo, che porta con sé elementi di novità e innovazione, ma anche

[217] C. SHEPARD, "How AI Overviews Shift Traffic from Publishers to Google", *Zyppy List*, May 8, 2025, https://zyppy.com/list/ai-overviews-to-google-ads/.

possibili effetti, in particolare per gli editori. Se, da un lato, la tecnologia offre un accesso rapido e diretto alle informazioni, dall'altro, l'IA può avere riverberi sulla sostenibilità economica dei giornali e dei *media online*, in quanto può ridurre il traffico *web* e, di conseguenza, le entrate pubblicitarie.

È una situazione, secondo tale analisi, che mette in evidenza alcune ricadute sulla concorrenza: **auto-preferenziazione (c.d. *self-preferencing*)**, in quanto potrebbero essere favoriti i propri servizi o quelli di partner selezionati, limitando l'accesso a contenuti di editori indipendenti; **utilizzo gratuito dei contenuti**, dato che le informazioni utilizzate per allenare l'IA trovano in gran parte origine da giornali e siti *web* senza alcuna compensazione per i produttori di questi contenuti; infine, **omogeneizzazione dell'informazione**, in quanto le risposte dell'IA potrebbero risultare uniformi e ridurre la pluralità delle opinioni e delle fonti, con gravi ripercussioni sul giornalismo e sulla democrazia dell'informazione.

In questo contesto, il ***Digital Markets Act* (DMA)**, approvato di recente in Europa, potrebbe rivelarsi uno strumento

fondamentale per regolare le problematiche legate all'IA. Ad esempio, le disposizioni di tale Atto relative al ***self-preferencing*** (articolo 6.5) e al **diritto degli editori** di richiedere un accesso equo e persino una compensazione per l'utilizzo dei loro contenuti (articolo 6.12) potrebbero intervenire efficacemente per arginare le menzionate pratiche anti-concorrenziali nel mercato dell'IA. Tuttavia, un punto al quale dedicare attenzione riguarda la definizione stessa dell'***AI Overview***: se questo servizio debba essere considerato come un servizio separato, soggetto alle normative del DMA, o come una funzione accessoria del motore di ricerca. Questa distinzione ha implicazioni significative per l'applicazione delle normative, dato che la sua classificazione influenzerebbe il modo in cui le piattaforme possono essere regolamentate e le pratiche concorrenziali monitorate.

Infine, un aspetto che merita particolare attenzione riguarda **il rischio di formazione di cartelli o pratiche collusive tra le grandi piattaforme tecnologiche**, così come nel settore del lavoro. Negli ultimi anni, le **autorità di concorrenza** hanno concentrato l'attenzione su un tipo di pratica anticoncorrenziale

che sta emergendo anche nel mercato del lavoro: i ***no-poaching agreements***[218]. Questi accordi prevedono che due o più aziende si impegnino a non assumere i dipendenti l'una dell'altra, limitando così la concorrenza per le risorse umane e impedendo la libera circolazione dei lavoratori tra le aziende, con un impatto negativo sul mercato del lavoro. Queste ipotesi risultano interessanti in quanto le implicazioni della digitalizzazione sono rilevanti anche nel contesto del mondo del lavoro che, proprio anche grazie alla digitalizzazione medesima e all'emergere di piattaforme tecnologiche, cambia ed evolve.

Un esempio che ha attirato l'attenzione della Commissione Europea è il caso delle aziende Glovo e Delivery Hero[219], nel quale si sospetta che le due società abbiano concordato di non rubare (*no-poaching*) i rispettivi dipendenti, creando una divisione del mercato del lavoro in cui i lavoratori non potevano

[218] D.J. POLDEN, "No-Poach Agreements: An overview of US, EU, and national case law", February 1 2024, articolo sul sito de Concurrences Antitrust Publications & Events, https://www.concurrences.com/en/bulletin/special-issues/no-poach-agreements/no-poach-agreements-an-overview-of-us-eu-and-national-case-law?id_rubrique=4388.

[219] EUROPEAN COMMISSION, "Commission opens investigation into possible anticompetitive agreements in the online food delivery sector", Press Release, July 23, 2024, https://ec.europa.eu/commission/presscorner/detail/en/ip_24_3908.

transitare liberamente tra le due piattaforme. La Commissione ha preparato accuse formali verso queste aziende per pratiche anticoncorrenziali, legate a collusioni sui salari e alla selezione geografica dei mercati in cui operano. Pratiche ritenute facilitate dalla partecipazione azionaria di Delivery Hero in Glovo, che ha permesso una coordinazione più stretta tra le due società.

Con la decisione del 2 giugno 2025, la Commissione europea ha accertato che, tra il 2018 e il 2022, due tra i principali operatori nel mercato del *food delivery*, Delivery Hero e Glovo, si sono coordinati al fine di non sottrarsi reciprocamente i dipendenti, condividere informazioni commercialmente sensibili e ripartirsi i mercati geografici all'interno dell'Unione[220].

Questo caso segna la **prima indagine della Commissione Europea sui cartelli nei contratti di non concorrenza** e si inserisce in un quadro più ampio di sforzi per garantire un

[220] EUROPEAN COMMISSION, "Commission fines Delivery Hero and Glovo €329 million for participation in online food delivery cartel", Press release, Jun 2, 2025, https://ec.europa.eu/commission/presscorner/detail/en/ip_25_1356. Cfr. anche il commento, "Commissione UE, sanzione "storica" da 329 milioni di euro a Delivery Hero e Glovo per il cartello nel mercato del *food delivery*", *Eurojus.it rivista*, giugno 2025, https://rivista.eurojus.it/commissione-ue-sanzione-storica-da-329-milioni-di-euro-a-delivery-hero-e-glovo-per-il-cartello-nel-mercato-del-food-delivery/.

mercato del lavoro equo e competitivo, dove le imprese non possano colludere per limitare le opportunità professionali dei lavoratori. Indagine che sottolinea l'importanza di monitorare non solo le pratiche anticoncorrenziali tra le imprese in relazione ai prodotti o servizi, ma anche l'impatto di tali pratiche sul mercato del lavoro. L'adozione di accordi di non concorrenza tra le aziende può minare la concorrenza per il talento e ridurre le opportunità di crescita professionale per i lavoratori, con conseguenze a lungo termine anche per l'innovazione e la competitività complessiva del settore.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

V.

POLITICHE GLOBALI E QUADRO NORMATIVO

Con il contributo di

Fabiana Di Porto, Marco Fontana, Andrea Bertolini, Giovanni Zarra

V.1 REGOLAMENTAZIONE DELL'INTELLIGENZA ARTIFICIALE A LIVELLO INTERNAZIONALE

Il quadro politico e regolatorio dell'intelligenza artificiale a livello internazionale è in continua evoluzione, anche in ragione delle innovazioni tecnologiche del settore e degli avvicendamenti politici nei governi.

V.1.1 UNIONE EUROPEA

A livello europeo, il 2024 è stato segnato dall'approvazione definitiva ed emanazione dell'AI Act, il Regolamento (UE) 2024/1689, pubblicato nella Gazzetta ufficiale europea il 12 luglio 2024[221] ed entrato in vigore il 1° agosto, anche se le norme in esso contenute troveranno applicazione gradualmente (v. *infra*).

L'AI Act rimane sinora il più importante esempio a livello internazionale di regolamentazione dell'intelligenza artificiale secondo un approccio "orizzontale": l'intelligenza artificiale è intesa come ambito di regolazione generale e autonomo, senza distinzioni basate sulle tecnologie specifiche impiegate. Tale approccio consente - nelle intenzioni del legislatore europeo - di adattarsi al futuro (inevitabile e rapido) sviluppo delle tecnologie che possono essere ricondotte al termine "intelligenza artificiale".

La scelta delle istituzioni europee muove probabilmente da considerazioni geopolitiche: con l'AI Act l'UE aspira ad adottare,

[221] Per esteso: Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n, 300/2008, (UE) n, 167/2013, (UE) n, 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale) (Testo rilevante ai fini del SEE), https://eur-lex.europa.eu/eli/reg/2024/1689/oj

per la prima volta, uno strumento regolatorio dedicato all'intelligenza artificiale "in quanto tale", che possa essere preso come riferimento e modello anche da altri ordinamenti, secondo il cd. "effetto Bruxelles"[222].

Inoltre, come espressamente affermato dalla Commissione europea[223], l'approccio orizzontale dell'AI Act è ispirato dal principio della neutralità tecnologica, uno dei principi che guida tradizionalmente le iniziative regolatorie dell'UE in ambito digitale e non solo.

L'approccio orizzontale però, in concreto, rischia di scontrarsi con l'ambiguità di fondo della nozione di "intelligenza artificiale", un termine "ombrello" dai confini incerti, che si

[222] V. sul tema, ad esempio, M. CARTA, "Il Regolamento UE sull'Intelligenza Artificiale: alcune questioni aperte", *Eurojus.it*, 2, 2024, in part. p. 190, in rete: https://rivista.eurojus.it/wp-content/uploads/pdf/Intelligenza-Artificiale-Carta-def-def.pdf. E anche G. FINOCCHIARO, *Diritto di internet*, Zanichelli Torino, IV ed., 2023, pp. 175-176; C. SIEGMANN, M. ANDERLJUNG, *The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market*, Centre for the Governance of the AI, August 2022, https://cdn.governance.ai/Brussels_Effect_GovAI.pdf. V. anche, in senso più critico, M. ALMADA, A. RADU, "The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy", *German Law Journal*, 25(4), 2024, pp.1-18, http://dx.doi.org/10.1017/glj.2023.108.

[223] Cfr. Comunicazione della Commissione, Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (Legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, Bruxelles, 21.4.2021, COM(2021) 206 final, par. 5.2.1.

traducono nella difficoltà di definire l'ambito di applicazione delle singole previsioni dell'atto (cfr. *infra*, *Box* 4.).

Un altro elemento fondamentale che connota l'AI Act è la finalità di protezione dei diritti fondamentali, la quale è evidenziata nel testo normativo fin dalle sue premesse iniziali. L'AI Act fonda l'applicazione degli obblighi previsti su differenti livelli di rischio, i quali hanno come riferimento principale il rischio di violazione dei diritti fondamentali, a cominciare dal diritto alla protezione dei dati personali (per maggiori dettagli si veda *infra* § V.2).

Tuttavia, l'AI Act ha anche l'obiettivo di promuovere lo sviluppo di tecnologie di intelligenza artificiale, attraverso l'armonizzazione delle legislazioni e l'eliminazione delle barriere che impediscono lo sviluppo di strumenti nel mercato comune. Perciò, le previsioni dell'AI Act richiamano nell'impostazione, quando non attraverso rinvii espliciti, la regolazione sulla sicurezza dei prodotti nel mercato interno europeo e i relativi processi di certificazione e vigilanza.

Questa impostazione ambivalente non è immune da potenziali criticità, da un lato, rispetto ad un eccesso di regolazione che possa frenare lo sviluppo di tecnologie nel contesto europeo[224] ; dall'altro lato, ci si interroga sulla reale capacità delle regole di tutelare i diritti fondamentali della persona. Con riguardo a quest'ultimo punto, esistono timori sull'eventualità che le regole dell'AI Act depotenzino il livello di tutela raggiunto e presidiato dall'applicazione del Regolamento UE 2016/679: il Regolamento generale sulla protezione dei dati personali (GDPR)[225].

Alcuni aspetti dell'AI Act (e, più in generale, dello sviluppo e utilizzo dei sistemi di IA) possono infatti risultare non in linea con i principi sul trattamento dei dati personali previsti dal GDPR e i relativi diritti fondamentali[226]: ad esempio, secondo tale

[224] Sul tema v., ad esempio, T. SCHREPEL, "Assessing the Impact of the European AI Act on Innovation Dynamics: Insights from Artificial Intelligences", *Journal of Competition Law & Economics*, 2025, nhaf007 (https://doi.org/10.1093/joclec/nhaf007); E. SCHNEIDER, "IA Act e i sistemi di rischio: lungimiranza o nostalgia?", Osservatorio sull'IA Act, IRPA, 12 dicembre 2024, https://www.irpa.eu/ia-act-e-i-sistemi-di-rischio-lungimiranza-o-nostalgia/.

[225] A tal proposito v. quanto osservato dal Comitato europeo per la protezione dei dati e dal Garante europeo per la protezione dei dati nel corso dell'*iter* di approvazione dell'AI Act: EDPB-GEPD, Parere congiunto 5/2021 sulla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale), 18 giugno 2021, https://www.edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_it.

[226] Per un inquadramento sistematico e generale, v. C. NOVELLI *ET AL.*, "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity", *Computer Law & Security Review*, vol. 55, 2024, https://doi.org/10.1016/j.clsr.2024.106066.

regolamento i dati devono essere trattati per finalità specifiche e per un tempo limitato; mentre, ai fini dell'addestramento di un modello di intelligenza artificiale, la finalità del trattamento e i termini di conservazione possono risultare difficilmente definibili a priori.

Un altro elemento critico per la protezione dei dati personali può derivare dalle disfunzioni di modelli di IA generativa come i *large language models*: ad esempio, le "allucinazioni" cui tali modelli sono soggetti possono generare informazioni errate su persone fisiche, in violazione del principio di esattezza dei dati previsto dal GDPR; le modalità di addestramento dei modelli possono portare a rivelare dati personali provenienti dai *corpora* utilizzati per il loro sviluppo o durante precedenti interazioni, anche attraverso pratiche malevole di *reverse engineering*, ponendo quindi problemi rispetto al diritto all'oblio. Le cd. *privacy enhancing technologies* (come ad esempio il cd. *machine unlearning* per eliminare dati specifici dalla "memoria" del modello, come ricordato *supra* § IV.3) possono mitigare tali rischi, ma difficilmente eliminarli del tutto.

Sul tema della compatibilità tra regole del GDPR e sviluppo di sistemi di IA, un contributo importante nell'ultimo periodo è arrivato dal parere del Comitato europeo per la protezione dei dati personali (EDPB) sull'uso dei dati personali per lo sviluppo e la diffusione di modelli di IA (v. *infra*, *Box* 1.).

Più in generale, l'AI Act è una componente fondamentale del programma di regolazione della società digitale in corso di attuazione a livello europeo, che ha visto negli ultimi anni l'emanazione di altri strumenti importanti, quali il Digital Services Act[227], il Digital Markets Act[228], il Data Act (su cui v. *infra*)[229] e il Data Governance Act[230], ma anche la nuova direttiva sulla cybersicurezza (NIS 2)[231] o il nuovo regolamento

[227] Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali), http://data.europa.eu/eli/reg/2022/2065/oj.

[228] Regolamento (UE) 2022/1925 del Parlamento europeo e del Consiglio del 14 settembre 2022 relativo a mercati equi e contendibili nel settore digitale e che modifica le direttive (UE) 2019/1937 e (UE) 2020/1828 (regolamento sui mercati digitali), https://eur-lex.europa.eu/eli/reg/2022/1925/2022-10-12.

[229] Regolamento (UE) 2023/2854 del Parlamento europeo e del Consiglio del 13 dicembre 2023, riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo e che modifica il regolamento (UE) 2017/2394 e la direttiva (UE) 2020/1828 (regolamento sui dati), https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=OJ%3AL_202302854.

[230] Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio del 30 maggio 2022 relativo alla *governance* europea dei dati e che modifica il regolamento (UE) 2018/1724 (Regolamento sulla *governance* dei dati), http://data.europa.eu/eli/reg/2022/868/oj.

[231] Direttiva (UE) 2022/2555 del Parlamento europeo e del Consiglio del 14 dicembre 2022 relativa a misure per un livello comune elevato di cibersicurezza nell'Unione, recante

sull'identità digitale (eIDAS 2)[232]. Questi interventi di regolazione a livello europeo riguardano evidentemente settori diversi, però contribuiscono alla definizione di uno "spazio unico digitale europeo" con regole comuni e ispirato alla tutela di diritti e principi fondamentali. I quali si applicano, peraltro, indipendentemente dal rapporto contrattuale in essere tra operatori e utenti di servizi digitali, purché i destinatari di questi ultimi siano collocati od operino nell'Unione.

In questo senso, gli strumenti normativi descritti esprimono anche la dimensione della sovranità (digitale) dell'Unione europea e possono essere considerate di natura costituzionale[233]. Si tratta di una costituzione "dal basso", cioè attraverso la regolamentazione settoriale dei rapporti tra operatori e utenti, non un esplicito e unico processo costituente. Ciò comporta la necessità di un

modifica del regolamento (UE) n. 910/2014 e della direttiva (UE) 2018/1972 e che abroga la direttiva (UE) 2016/1148 (direttiva NIS 2), http://data.europa.eu/eli/dir/2022/2555/oj.

[232] Regolamento (UE) 2024/1183 del Parlamento europeo e del Consiglio, dell'11 aprile 2024, che modifica il regolamento (UE) n. 910/2014 per quanto riguarda l'istituzione del quadro europeo relativo a un'identità digitale, si veda: http://data.europa.eu/eli/reg/2024/1183/oj.

[233] Si veda sul tema, ad esempio, F. PIZZETTI, "Da Tallin all'AI Act, così l'Ue costruisce la sua Costituzione digitale", *AgendaDigitale.EU*, 7 maggio 2024, https://www.agendadigitale.eu/cultura-digitale/pizzetti-da-tallin-allai-act-cosi-lue-costruisce-la-sua-costituzione-digitale/.

coordinamento delle disposizioni anche al fine di evitare rischi di sovrapposizione e incoerenze tra i singoli testi normativi.

BOX 1. PARERE EDPB DEL 18 DICEMBRE 2024 SU USO DEI DATI PERSONALI PER LO SVILUPPO E LA DIFFUSIONE DI MODELLI DI IA.[234]

A inizio settembre 2024, l'Autorità per la protezione dei dati irlandese ha richiesto un parere al Comitato europeo per la protezione dei dati personali (EDPB) ai sensi dell'art. 64, paragrafo 2 del GDPR, con riguardo al trattamento di dati personali nell'ambito dello sviluppo e dell'impiego di modelli di intelligenza artificiale (IA).

Il parere richiesto riguarda quattro quesiti fondamentali: (1.) quando e come i modelli di IA possono essere considerati anonimi (primo quesito); (2.) come i titolari del trattamento possono dimostrare l'utilizzo dell'interesse legittimo quale base giuridica per il trattamento dei dati personali nella fase di sviluppo o (3.) di utilizzo del modello; (4.) quali conseguenze ha rispetto all'utilizzo del modello il trattamento illecito di dati personali nella fase di sviluppo del modello.

Per quanto riguarda il primo quesito, l'EDPB indica che devono essere le autorità nazionali a stabilire caso per caso se un modello può essere considerato anonimo. Questo può dirsi tale quando la probabilità di estrarre direttamente dai modelli dati personali o ottenere gli stessi (anche non intenzionalmente) da richieste ai modelli sia insignificante. L'anonimato deve essere valutato dalle autorità competenti sulla base della documentazione

[234] EDPB, Parere 28/2024 su taluni aspetti relativi alla protezione dei dati ai fini del trattamento dei dati personali nel contesto dei modelli di IA, 18 dicembre 2024, https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_it.

fornita dagli sviluppatori a comprova delle loro garanzie di anonimato (p. es. eliminazione di dati personali dal set di dati per l'addestramento).

Per quanto riguarda il secondo e terzo quesito, l'EDPB dà indicazioni riguardo al test che le autorità nazionali devono compiere nel valutare l'ammissibilità del legittimo interesse quale base giuridica del trattamento dei dati per lo sviluppo o nell'utilizzo dei modelli di IA.

A tale riguardo, il parere ripropone un modello con tre step: l'individuazione dell'interesse legittimo; la valutazione dell'effettiva necessità e adeguatezza del trattamento per il perseguimento dell'interesse (*necessity test*); l'esistenza di un interesse contrario in capo all'interessato che prevalga sull'interesse del titolare (*balancing test*). Il parere offre indicazioni dettagliate per ciascun passaggio del test, con esempi concreti su come considerare in particolare il bilanciamento tra interessi dell'individuo e l'interesse legittimo (p. es. verificare il contesto in cui il dato viene raccolto). Inoltre, suggerisce delle misure di mitigazione da adottare quando la posizione del singolo interessato è da considerarsi prevalente.

Per quanto riguarda il quarto quesito, l'EDPB ipotizza tre scenari diversi, ferma restando sempre la valutazione dell'autorità nazionale sui casi specifici. Nel primo scenario, è lo stesso titolare del trattamento nella fase dello sviluppo del modello (con trattamento illecito di dati personali) che impiega il modello stesso. In questo caso, occorre valutare in concreto se l'impiego del modello abbia una finalità diversa rispetto allo sviluppo e, quindi, se l'illecito trattamento nella prima fase si ripercuota necessariamente o meno anche sulla seconda.

Nel secondo scenario, il titolare del trattamento dei dati per l'impiego del modello è diverso dal titolare del trattamento (illecito) per il suo sviluppo. In questo caso, occorre valutare se il *deployer* abbia approfondito adeguatamente

la liceità del trattamento nella fase di sviluppo, in particolare quando quest'ultima sia stata oggetto di indagini o accertamenti ufficiali e l'utilizzo del modello comprenda rischi rilevanti.

Infine, il terzo scenario riguarda il caso in cui i dati trattati illecitamente in fase di sviluppo vengano successivamente anonimizzati per la fase di impiego. In tal caso, non ci sono ripercussioni dirette sulla liceità del trattamento nell'impiego del modello, anche se il titolare del trattamento non cambia. Se l'utilizzo non implica il trattamento di nuovi dati personali, il GDPR non troverà applicazione; se invece saranno trattati nuovi dati, la liceità del trattamento va valutata indipendentemente da quanto accaduto in fase di sviluppo.

Uno strumento utile per una panoramica sui programmi in materia di IA della Commissione von der Leyen, entrata in carica ufficialmente il 1° dicembre 2024, sono le "Lettere di missione" per i Commissari. Con queste ultime, la Presidente von der Leyen assegna compiti e portafogli ai nuovi Commissari fino al 2029. Per tale arco temporale la Commissaria più coinvolta nelle politiche sull'IA è Henna Virkkunen, Vicepresidente esecutiva "per la sovranità tecnologica, la sicurezza e la democrazia".

Gli ambiti in cui opererà la Commissione sono, ad esempio:

a.) dare impulso all'innovazione dell'IA nei primi cento giorni della Commissione europea, in primo luogo fornendo alle

start-up l'accesso ai supercomputer nell'ambito della strategia "Fabbriche di IA"[235];

b.) accelerare l'adozione e la messa in opera delle tecnologie di IA nell'industria e nel settore pubblico;

c.) istituire un Consiglio di ricerca dell'UE sull'IA per coordinare e far progredire la ricerca sull'IA in Europa;

d.) sviluppare una strategia sull'uso delle tecnologie digitali, compresa l'IA, per rendere i sistemi giudiziari civili e penali dell'Unione europea più efficienti, resilienti e sicuri;

e.) elaborare una proposta di legge dell'UE per lo sviluppo del *cloud* e dell'IA per aumentare la capacità di calcolo, come evidenziato nel rapporto Draghi[236] sulla competitività dell'UE. Inoltre, creare un quadro a livello dell'Unione per fornire "capitale computazionale" alle PMI innovative.

Oltre all'attenzione sull'intelligenza artificiale, le Lettere di missione contengono molti riferimenti alla sicurezza

[235] Cfr. COMMISSIONE EUROPEA, AI Factories, https://digital-strategy.ec.europa.eu/it/policies/ai-factories e la Lettera di missione per la Commissaria Henna Virkkunen, disponibile alla pagina: https://commission.europa.eu/document/3b537594-9264-4249-a912-5b102b7b49a3_en.. Su questa lettera cfr. *infra* e vedi anche nota 239.

[236] EUROPEAN COMMISSION, *The future of European competitiveness* (part. A and part B), September 2024, scaricabile dalla pagina: https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en.

informatica, la quale compare sia in un obiettivo generale (assicurare standard elevati di cybersicurezza), sia in obiettivi più specifici, quali la sicurezza informatica nel contesto sanitario e il miglioramento dei processi di adozione di schemi di certificazioni per la cybersicurezza.

Gli altri temi ricorrenti riguardano l'infrastruttura digitale. Ad esempio, le Lettere comprendono: lo sviluppo di una politica *cloud* unica a livello europeo per le pubbliche amministrazioni e gli appalti pubblici; un piano europeo a lungo termine per i chip quantistici (sempre su suggerimento del rapporto Draghi); una nuova legge sulle reti digitali (Digital Networks Act, DNA) per contribuire a potenziare la banda larga sicura ad alta velocità e per incentivare e incoraggiare gli investimenti nelle infrastrutture digitali; l'incremento dell'infrastruttura pubblica digitale per garantire che le imprese possano accelerare e semplificare le operazioni e ridurre i costi amministrativi (anche attraverso l'EU Digital Identity Wallet[237]).

[237] Ai sensi dell'art. 3, n, 42) del Regolamento eIDAS (Regolamento (Ue) 910/2014), come modificato dal recente Regolamento (Ue) n. 2024/1183, il *Digital Identity Wallet* ("Portafoglio europeo di identità digitale") è: "un mezzo di identificazione elettronica che consente all'utente di conservare, gestire e convalidare in modo sicuro dati di identità personale e

Infine, viene posta attenzione sui dati, con la previsione di una nuova Strategia UE per un'Unione dei Dati (prevista per il terzo trimestre 2025), che muova dalla Strategia dell'Unione per i dati del 2020[238] e dalle regole esistenti per coordinarle, creando un quadro legale semplificato, chiaro e coerente per la condivisione dei dati tra imprese e amministrazioni, rispettando al contempo standard elevati di *privacy* e sicurezza. Nello stesso filone possono essere ricondotti una indagine europea sull'impatto dei *social media* e la lotta a tecniche non etiche utilizzate *online*, come i *dark patterns*[239].

attestati elettronici di attributi al fine di fornirli alle parti facenti affidamento sulla certificazione e agli altri utenti dei portafogli europei di identità digitale, e di firmare mediante firme elettroniche qualificate o apporre sigilli mediante sigilli elettronici qualificati".

[238] Vedi la pagina della Commissione Europea sulla strategia europea per i dati: https://digital-strategy.ec.europa.eu/it/policies/strategy-data. La Strategia UE per i dati del 2020 (Commissione europea, Comunicazione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, Una strategia europea per i dati, Bruxelles, 19.2.2020 COM(2020) 66 final), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066.

[239] Cfr. La Lettera di missione per la Commissaria Henna Virkkunen, già citata alla nota 235 e vedi anche P. ECKHARDT *ET AL*., "The EU's Digital Ambitions. Candidates, Portfolios and EU Initiatives for the EU Commission 2024-2029", *Missions Letters 2/2024*, cepInput special, CEP - Centres for European Policy Network, 29 settembre 2024, https://www.cep.eu/eu-topics/details/mission-letters-digital-policy.html. Quest'ultima analisi elenca anche i Commissari coinvolti nell'attuazione delle politiche digitali, i quali, oltre alla Vicepresidente Virkkunen, possono essere individuati in Stéphane Séjourné (Vicepresidente esecutivo Prosperità e strategia industriale), Olivér Várhelyi (Salute e benessere degli animali), Michael McGrath (Democrazia, giustizia, Stato di diritto e tutela dei consumatori), Magnus Brunner (Affari interni e migrazione), Ekaterina Zaharieva (Start-up, ricerca e innovazione), Glenn Micallef (Equità intergenerazionale, giovani, cultura e sport).

La Commissione europea ha confermato il proprio impegno sullo sviluppo di tecnologie di intelligenza artificiale nel contesto europeo nell'ambito dell'AI Summit di Parigi (6-11 febbraio 2025)[240], con la previsione di un piano di investimenti di duecento miliardi di euro (inclusi i contributi privati) attraverso l'iniziativa InvestAI, inclusi venti miliardi di euro per lo sviluppo di infrastrutture industriali dedicate (*AI Giga Factories*)[241]. Ancora e anche sulla base del Rapporto Draghi, la Commissione ha inoltre emanato una comunicazione recante una "bussola sulla competitività europea"[242] di gennaio 2025, in cui si annunciano una serie di iniziative volte a promuovere lo sviluppo dell'intelligenza artificiale, in un'ottica di promozione della competitività europea.

In questa prospettiva, la Commissione europea nell'aprile 2025 ha pubblicato *l'AI Continent Action Plan* (che include il piano di

[240] Per una sintesi dei risultati del Summit cfr. il documento della Presidenza della Repubblica francese disponibile al link: https://www.elysee.fr/en/sommet-pour-l-action-sur-l-ia.

[241] Il comunicato relativo all'iniziativa "EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence" è disponibile qui: https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467.

[242] Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, A Competitiveness Compass for the EU, Brussels, 29.1.2025 COM(2025) 30 final, https://commission.europa.eu/document/download/10017eb1-4722-4333-add2-e0ed18105a34_en.

investimenti in super computer per l'IA) come uno degli elementi fondamentali per lo sviluppo dell'innovazione tecnologica nell'economia europea. **L'*AI Continent Action Plan* individua cinque ambiti principali per permettere al sistema europeo di guadagnare il terreno nel contesto dell'intelligenza artificiale**: infrastruttura computazionale, dati di alta qualità, promozione dello sviluppo di algoritmi di intelligenza artificiale e della loro adozione nei contesti industriali, sviluppo competenze in ambito IA e semplificazione regolatoria[243].

In tale contesto, la Commissione ha proposto l'adozione di un nuovo atto normativo per lo sviluppo dell'infrastruttura *cloud* e IA (Cloud and AI Development Act), con una "*call for evidence*" - e relativa consultazione - aperte fino a luglio 2025[244]. Sempre nell'ambito dell'*AI Continent Action Plan*, la Commissione ha anche pubblicato due *calls for evidence* e due consultazioni per

[243] Cfr. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, AI Continent Action Plan, Brussels, 9.4.2025 COM(2025) 165 final, scaricabile dalla pagina: https://digital-strategy.ec.europa.eu/en/library/ai-continent-action-plan; il testo del solo Piano è anche disponibile qui: https://commission.europa.eu/topics/eu-competitiveness/ai-continent_en.

[244] La pagina per la *call for evidence* è accessibile da qui: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/14628-AI-Continent-new-cloud-and-AI-development-act_en.

raccogliere input per l'emanazione di due strategie distinte, ossia l'*Apply AI Strategy*[245] e l'*European Strategy for AI in Science*[246].

V.1.2 STATI UNITI

L'avvicendamento dell'amministrazione U.S.A. ha segnato un cambio di indirizzo nelle politiche federali degli Stati Uniti sull'intelligenza artificiale.

Il nuovo Presidente ha revocato l'ordine esecutivo emanato dal suo predecessore (Executive Order 14110 of October 30, 2023 - *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*) il giorno stesso del suo insediamento, il 20 gennaio 2025[247]. E con un nuovo Ordine Esecutivo del 23 gennaio 2025[248], la Presidenza Trump ha introdotto tre azioni in materia di intelligenza artificiale e attività finanziarie digitali, indicando così una nuova

[245] La pagina per la *AI Strategy call for evidence* è accessibile da qui: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/14625-Apply-AI-Strategy-strengthening-the-AI-continent_en.

[246] La pagina per la *call for evidence* della *European Strategy for AI in Science* è accessibile da qui: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/14547-A-European-Strategy-for-AI-in-science-paving-the-way-for-a-European-AI-research-council_en.

[247] Executive Order 14148 of January 20, 2025, Initial Rescission of Harmful Executive Orders, https://www.federalregister.gov/documents/2025/01/28/2025-01901/initial-rescissions-of-harmful-executive-orders-and-actions.

[248] Executive Order 14179 of January 23, 2025, Removing Barriers to American Leadership in Artificial Intelligence, https://www.federalregister.gov/documents/2025/01/31/2025-02172/removing-barriers-to-american-leadership-in-artificial-intelligence.

direzione, che include anche la riorganizzazione del Consiglio dei consulenti del Presidente per la scienza e la tecnologia.

L'Ordine Esecutivo muove dalla constatazione del primato degli Stati Uniti nell'ambito dell'intelligenza artificiale, e sostiene che, per mantenere tale primato, sia necessario: "sviluppare sistemi IA che siano liberi da bias ideologici o programmi sociali 'ingegnerizzati'". A tal fine, la dichiarazione di *policy* dell'ordine riguarda il sostegno e la valorizzazione del "dominio globale in materia di IA" degli U.S.A., in modo da: "promuovere la prosperità umana, la competitività economica e la sicurezza nazionale".

Più concretamente, l'Ordine prevede un periodo di 180 giorni durante il quale lo staff della Casa Bianca (in particolare, l'Assistente del Presidente per la Scienza e Tecnologia, il Consigliere Speciale per l'IA e le Crypto, nonché l'Assistente del Presidente per Sicurezza Nazionale) deve predisporre un piano per attuare la nuova politica e assicurarsi che le agenzie federali prendano misure per annullare le direttive precedenti incompatibili con essa.

Non vi è dunque un obbligo generalizzato di superare le indicazioni dei precedenti ordini esecutivi, ma i piani di attuazione dello stesso dovranno individuare quali azioni siano di ostacolo allo sviluppo delle nuove politiche.

L'Ordine prevede anche di ridefinire le regole sull'uso dei sistemi di IA da parte del governo federale, in particolare sull'uso da parte di agenzie (*Memorandum* M-24-10 del 28 marzo 2024) e sull'acquisto responsabile dei sistemi IA (*Memorandum* M-24-18 del 24 settembre 2024). Il Direttore dell'Ufficio per la Gestione e il Budget ha sessanta giorni di tempo per riscrivere le regole e allinearle alla nuova politica.

A livello statale si possono rilevare diverse iniziative legislative che riguardano lo sviluppo e l'utilizzo di sistemi di IA. La più importante di queste, nel corso del 2024, è stata la proposta di legge dello Stato di California SB-1047[249]. La legislazione della California risulta particolarmente rilevante per l'importanza dell'economia dello Stato, il quale è sede - come noto - di gran

[249] SB-1047 Safe and Secure Innovation for Frontier Artificial Intelligence Models Act (2023-2024), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB1047.

parte delle società leader globali di mercato dell'intelligenza artificiale e del digitale più in generale. Il "*bill*" è stato approvato dal Parlamento statale a settembre 2024, ma successivamente oggetto di veto da parte del Governatore dello Stato, Gavin Newsom. La proposta non è dunque diventata legge.

Il SB-1047 conteneva diversi elementi ritenuti assimilabili all'approccio regolatorio proprio dell'AI Act europeo, a cominciare dalla volontà di regolare l'adozione di tecnologie di IA indipendentemente da impieghi o settori specifici e le tipologie di obblighi previsti (certificazioni, processi interni e controlli esterni). La differenza fondamentale rispetto all'AI Act risiede, invece, nell'ambito di applicazione del progetto di legge californiano, che riguarda(va) soltanto i modelli di IA (non quindi i "sistemi" basati sui modelli) e soltanto di dimensioni davvero molto grandi, definite sulla base degli investimenti e la potenza di calcolo necessari per il loro sviluppo. Inoltre, gli obblighi previsti dal SB-1047 riguarda(va)no soltanto lo sviluppatore (o "fornitore", secondo la terminologia europea) e non il *deployer* dei modelli.

Alcune differenze più specifiche della proposta californiana rispetto al modello europeo riguardano le misure tecniche a carico degli sviluppatori dei modelli, che comprendono: (1.) la possibilità di "spegnere" completamente e immediatamente ("*full shutdown*") il processo di addestramento del modello in caso di rischio di danni critici, (2.) gli obblighi dei fornitori di potenza di calcolo, nel caso in cui la potenza fornita a un loro cliente sia tale da poter addestrare un modello incluso nell'ambito di applicazione delle norme, (3,) il potere dell'*Attorney General* statale di agire in giudizio per il risarcimento del danno causato dai modelli - con un limite del 10% dei costi per l'utilizzo della potenza computazionale necessaria ad addestrare i modelli - o per ottenere altri rimedi (p. es. inibitoria).

La proposta, come detto, è stata respinta tramite veto dal Governatore statale[250], con una motivazione duplice: da un lato, la limitazione dell'applicazione delle regole ai soli modelli di grande dimensioni avrebbe rischiato - secondo la nota accompagnatoria del veto - di generare un "falso senso di sicurezza" nel pubblico;

[250] SB-1047 Safe and Secure, *ibidem*.

dall'altro, la mancanza di una distinzione dei casi di applicazione dei modelli avrebbe portato a standard troppo stringenti "anche per le funzioni più basilari", ostacolando così lo sviluppo di modelli in ambiti utili e con pochi rischi.

Il promotore della norma (Senatore Wiener) ha proposto, ad inizio 2025, un nuovo progetto di legge che si pone come obiettivo la regolazione dello sviluppo in senso responsabile di sistemi di intelligenza artificiale su larga scala[251].

Esistono comunque altre iniziative a livello statale che si sono tradotte in legge: la stessa California ha approvato, ad esempio, nel corso del 2024 una legge (*Chapter* 817[252]), che prevede l'obbligo dal 1° gennaio 2026 per gli sviluppatori di sistemi o servizi di IA generativa di pubblicare informazioni relative ai dati impiegati nello sviluppo dei sistemi o servizi stessi.

Lo Stato del Colorado, sempre a titolo esemplificativo, ha invece approvato nel corso del 2024 una legge statale riguardante la

[251] California Legislature, SB-53 CalCompute: foundation models: whistleblowers (2025-2026), https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202520260SB53.

[252] California Legislature, AB-2013, Generative artificial intelligence: training data transparency, https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202320240AB2013.

"protezione dei consumatori nell'interazione con i sistemi di intelligenza artificiale"[253], i cui obblighi principali - come per la legge californiana - entreranno in vigore nel (febbraio) 2026. La legge ha tratti in comune con il modello dell'AI Act, anche se si applica solo nel contesto di decisioni che abbiano un impatto sui consumatori: distingue tra livelli di rischio e prevede obblighi specifici per sviluppatori e "deployer" nel contesto dei sistemi ad alto rischio.

Da ultimo, lo Stato dello Utah ha approvato una legge[254] che prevede obblighi di trasparenza nell'impiego di sistemi di intelligenza artificiale generativa, a carico di sviluppatori e deployer, anche questa volta nell'ambito del diritto per la protezione dei consumatori.

V.1.3 CINA

L'ultimo anno è stato caratterizzato da novità in ambito IA anche nell'ordinamento cinese. Nel 2024, la Cina ha intensificato gli

[253] Colorado General Assembly, SB24-205, Consumer Protections for Artificial Intelligence, https://leg.colorado.gov/bills/sb24-205.

[254] Utah State Legislature, S.B. 149 Artificial Intelligence Amendments, https://le.utah.gov/~2024/bills/static/SB0149.html.

sforzi per stabilire un quadro normativo globale per essa. Un impegno che si è concretizzato nella partecipazione attiva a iniziative internazionali di *governance* dell'IA, mirate a creare standard condivisi per garantire lo sviluppo sicuro ed etico delle tecnologie emergenti.

In particolare, in attuazione della "Iniziativa per la *governance* globale dell'AI", lanciata dal Presidente cinese nell'ottobre del 2023[255], il Governo cinese (segnatamente, il Comitato Tecnico Nazionale 260 sulla cybersicurezza) ha pubblicato l'*AI Safety Governance Framework*[256], un documento (in inglese) che si basa su principi chiave come sicurezza inclusiva e sostenibile, prevenzione proattiva dei rischi e cooperazione internazionale. Il *Framework* non ha un valore vincolante, però è considerato come un'indicazione valida sia quale *best practice* per il settore (nel contesto cinese, ma anche internazionale), nonché come un documento di indirizzo per le politiche governative e le possibili ulteriori iniziative legislative in materia.

[255] Global AI Governance Initiative, testo in inglese disponibile qui: http://no.china-embassy.gov.cn/eng/lcbt/lcwj/202401/P020240112008151194499.pdf

[256] National Technical Committee 260 on Cybersecurity of Standardization Administration of China 2024.9, https://www.tc260.org.cn/upload/2024-09-09/1725849192841090989.pdf.

Anche il *Framework*, come l'AI Act europeo, ha un approccio *risk-based*. A differenza del Regolamento europeo, però, il *Framework* non si basa su una classificazione di livelli di gravità del rischio bensì sull'origine e l'area di riferimento del rischio. In questo senso, propone due macrocategorie:

1. rischi intrinseci della tecnologia, a loro volta suddivisi in a.) rischi da modelli e algoritmi; b.) rischi da dati e c.) rischi da sistemi di IA;
2. rischi applicativi, a loro volta suddivisi in rischi nel cyberspazio, rischi nel mondo reale, rischi cognitivi e rischi etici.

Il *Framework* contiene poi delle raccomandazioni sulle misure di mitigazione da adottare sia dal punto di vista tecnologico sia dal punto di vista regolatorio. In quest'ultimo contesto, il documento include suggerimenti che richiamano il modello regolatorio europeo, con la proposta di creare categorie di rischio basate su caratteristiche, funzioni e ambiti applicativi dei sistemi di IA, nonché la registrazione e l'applicazione di obblighi specifici per i sistemi più rischiosi.

Da ultimo, il *Framework* include linee guida sullo sviluppo e l'utilizzo dei sistemi di IA per sviluppatori, fornitori di servizi e utenti, con enfasi sulla trasparenza algoritmica, il miglioramento della robustezza dei sistemi e il rafforzamento delle infrastrutture di sicurezza.

L'approccio "orizzontale" del *Framework* (che si applica di per sé all'ambito dei sistemi di IA senza ulteriori precisazioni) si ritrova anche in una proposta di "Legge sull'intelligenza artificiale della Repubblica Popolare Cinese"[257], redatta da un gruppo di giuristi e pubblicata (anche in inglese) nella prima metà del 2024. Quest'ultima - che richiama come detto alcuni aspetti del *Framework* - non risulta però essere attualmente in corso di approvazione dagli organi legislativi della Repubblica Popolare.

Il sistema cinese conosce invece già delle misure regolatorie più settoriali, adottate nel corso dell'ultimo quinquennio, quali:

[257] Artificial Intelligence Law of the People's Republic of China (Draft for Suggestions from Scholars), per la traduzione in inglese, a cura del Center for Security and emerging Technology, v. https://cset.georgetown.edu/wp-content/uploads/t0592_china_ai_law_draft_EN.pdf

· le misure provvisorie per la gestione dei servizi di intelligenza artificiale generativa, entrate in vigore il 15 agosto 2023[258];

· le disposizioni amministrative sulla "sintesi profonda" (*deep synthesis* - produzione di contenuti tramite AI generativa/*deep learning*) nei servizi di informazione basati su internet, entrate in vigore il 10 gennaio 2023259;

· le disposizioni amministrative sugli algoritmi di raccomandazione nei servizi di informazione basati su Internet, entrate in vigore il 1° marzo 2022[260].

Il panorama regolatorio prevede dunque già misure governative, che impongono obblighi in relazione allo sviluppo e all'impiego di sistemi di IA, a volte anche piuttosto rilevanti. Un esempio in questo senso è l'obbligo di registrazione presso l'autorità centrale competente (cd. "*Cyber Administration of China*") degli algoritmi di raccomandazione che possano influenzare l'opinione pubblica.

[258] Sintesi in inglese e *link* all'originale in cinese disponibile sul sito della Library of Congress americana, "China: Generative AI Measures Finalized", https://www.loc.gov/item/global-legal-monitor/2023-07-18/china-generative-ai-measures-finalized/.

[259] "China: Provisions on Deep Synthesis Technology Enter into Effect", è una sintesi in inglese con *link* alla versione originale in cinese, sul sito della *Library of Congress*, https://www.loc.gov/item/global-legal-monitor/2023-04-25/china-provisions-on-deep-synthesis-technology-enter-into-effect/.

[260] Per il testo originale cfr. https://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm.

L'Autorità ha approvato diversi "lotti" di algoritmi in attuazione di tale misura anche nel corso del 2024[261].

Da una diversa prospettiva, l'ultimo anno ha segnato anche un'accelerazione nella capacità di sviluppo di strumenti di intelligenza artificiale innovativi anche per il contesto globale da parte del sistema cinese. Almeno dal 2017, anno di pubblicazione del "Piano di Sviluppo dell'Intelligenza Artificiale di Nuova Generazione"[262], il Governo cinese ha dato una particolare priorità alla strategia a lungo termine sull'intelligenza artificiale, con l'obiettivo per la Cina di diventare leader nel settore entro il 2030. Il piano prevedeva il raggiungimento di un livello di primato mondiale per le tecnologie e le applicazioni dell'industria cinese entro il 2025. A tal fine, il Governo cinese ha investito molto nella promozione politica e finanziaria del settore, così come nella formazione (in Cina o all'estero) del personale necessario per lo sviluppo dell'industria.

[261] "China: CAC Domestic Deep Synthesis Service Algorithm Filing List, Digital Policy Alert", sono disponibili aggiornamenti in lingua inglese e rimandi alle pubblicazioni ufficiali, https://digitalpolicyalert.org/change/6175.

[262] Per una traduzione in inglese del documento cfr. "China's 'New Generation Artificial Intelligence Development Plan" (2017), DigiChina, Stanford University, https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

Il risultato più appariscente di questa politica è stata la diffusione, tra fine 2024 e inizio 2025, dei modelli di intelligenza artificiale generativa sviluppati dalla società cinese DeepSeek. Tali modelli - e in particolare i più recenti DeepSeek-V3, DeepSeek-R1-Zero e DeepSeek-R1 - hanno mostrato livelli di *performance*, anche in relazione a capacità di "ragionamento", paragonabili ai *large language model* prodotti dai maggiori player di mercato statunitensi, con modalità, a detta degli sviluppatori, più efficienti in termini di parametri necessari per l'addestramento e di capacità computazionale impiegata.

BOX 2. DICHIARAZIONE SUL FUTURO DI INTERNET: CONVERGENZE E DIVERGENZE TRA LE GRANDI POTENZE.

Sono passati quasi tre anni dalla proposta presentata nel 2022 - su iniziativa degli Stati Uniti, dell'Unione europea e di altri Stati partner internazionali[263]- di una Dichiarazione per il Futuro di Internet[264], cui hanno aderito, nei mesi

[263] La Dichiarazione è stata presentata dalla Commissione europea il 28 aprile 2022. Gli Stati partner all'epoca, oltre a Stati Uniti, Unione europea e Stati membri UE, erano Albania, Andorra, Argentina, Australia, Capo Verde, Canada, Colombia, Costa Rica, Repubblica Dominicana, Georgia, Islanda, Israele, Giamaica, Giappone, Kenya, Kosovo, Maldive, Isole Marshall, Micronesia, Moldavia, Montenegro, Nuova Zelanda, Niger, Macedonia del Nord, Perù, Regno Unito, Serbia, Taiwan, Trinidad e Tobago, Ucraina e Uruguay. Cfr. il comunicato della Commissione alla pagina: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2695

[264] Si veda, sul sito del Dipartimento di Stato USA: *Declaration for the Future of the Internet*, https://www.state.gov/declaration-for-the-future-of-the-internet.

successivi, settanta firmatari[265]. La Dichiarazione si fonda su alcuni principi chiave, volti a garantire un internet aperto, libero e sicuro per tutti.

Tra i principi fondamentali vi è la protezione dei diritti umani e delle libertà fondamentali, nel riconoscere l'importanza di un ecosistema digitale che salvaguardi la *privacy* e la libertà di espressione. La dichiarazione promuove inoltre un internet globale, favorendo il libero flusso di informazioni senza restrizioni ingiustificate, garantendo accesso equo e conveniente alle infrastrutture digitali per i cittadini, indipendentemente dal loro contesto socioeconomico.

Un altro aspetto cruciale è la necessità di rafforzare la fiducia nel sistema digitale attraverso misure che tutelino la *privacy* e la sicurezza *online*. Infine, viene ribadita l'importanza di una *governance* multi-stakeholder, che coinvolga governi, settore privato, società civile e comunità tecnica nella gestione e nello sviluppo di internet.

Tuttavia, nonostante questi intenti condivisi sulla carta, la realtà dei fatti mostra una tendenza divergente. Molti dei governi di Stati leader dal punto di vista tecnologico a livello internazionale stanno adottando misure che vanno nella direzione opposta, favorendo la creazione di ecosistemi digitali chiusi e nazionalizzati.

La Cina, ad esempio, ha costruito una rete internet altamente regolata, con accesso limitato alle piattaforme occidentali come Google e Facebook, mentre i suoi servizi digitali, come il *cloud* di Tencent e i mini-program di WeChat, rimangono appannaggio esclusivo degli utenti cinesi[266]. La Russia ha

[265] L'elenco dei firmatari è reperibile qui: Declaration for the Future of Internet | Shaping Europe's digital future (ultimo aggiornamento 2 marzo 2023).

[266] Si veda sul tema, ad esempio, l'approfondimento di Stanford, *Free speech vs Maintaining Social Cohesion. A Closer Look at Different Policies*, FreeExpressionVsSocialCohesion/china_policy.html.

rafforzato la sua autonomia digitale, promuovendo piattaforme locali come VK, non particolarmente utilizzate al di fuori del paese[267].

Nel frattempo, gli Stati Uniti hanno vietato l'installazione di applicazioni come TikTok su dispositivi governativi[268] e hanno poi previsto la cessione dell'applicazione da parte della società proprietaria (Bytedance Ltd), pena il divieto di operare negli Stati Uniti[269], evidenziando come la frammentazione di internet sia un fenomeno in crescita su scala globale.

A questa tendenza si aggiunge un ulteriore problema: il controllo delle infrastrutture digitali è sempre più concentrato nelle mani di pochi attori privati, piuttosto che di Stati nazionali. Il settore del *cloud computing*, ad esempio, vede una crescente centralizzazione intorno ai cosiddetti *hyperscaler*, come Amazon Web Services, Google Cloud e Microsoft Azure, che detengono un oligopolio sulla gestione dei dati a livello globale. Una situazione che solleva interrogativi sulla sovranità digitale, in quanto le decisioni strategiche riguardanti internet e le sue infrastrutture potrebbero essere guidate più dagli interessi di colossi tecnologici privati che da quelli dei governi e dei cittadini.

In questo contesto, il futuro di internet appare incerto. Da un lato, si prospetta una rete sempre più frammentata, suddivisa in sfere d'influenza nazionali e regionali; dall'altro, si delinea uno scenario in cui pochi attori privati

[267] M. MEAKER, "Come il Cremlino ha preso il controllo di Vk, il *Facebook* russo", *Wired*, 12 giugno 2022, https://www.wired.it/article/russia-vkontakte-influenza-governo/.

[268] S.1143 - No TikTok on Government Devices Act, Act 117th Congress (2021-2022), https://www.congress.gov/bill/117th-congress/senate-bill/1143.

[269] Cfr. Protecting Americans from Foreign Adversary Controlled Applications Act (in sigla, PAFACA), https://www.congress.gov/bill/118th-congress/house-bill/7521, per un commento v. anche: https://www.law.cornell.edu/uscode/text/15/9901. Al termine dei 270 giorni previsti dalla legge per l'adeguamento della società alle previsioni, il Presidente ha prorogato l'entrata in vigore del divieto nei confronti di Bytedance per ulteriori 75 giorni (Cfr. Executive Order 14166 of January 20, 2025 Application of Protecting Americans From Foreign Adversary Controlled Applications Act to TikTok, https://www.federalregister.gov/documents/2025/01/30/2025-02087/application-of-protecting-americans-from-foreign-adversary-controlled-applications-act-to-tiktok).

detengono il controllo delle infrastrutture chiave, influenzando non solo l'economia digitale, ma anche la *governance* e la sicurezza globale. La vera sfida sarà conciliare gli ideali espressi nella Dichiarazione con le politiche concrete che gli Stati adotteranno nei prossimi anni, cercando di bilanciare sovranità nazionale, sicurezza e apertura della rete globale.

A ciò si aggiunge, sul piano normativo, la questione della differenza nella regolazione delle tecnologie digitali nei diversi sistemi: la produzione di nuove regole è continua e ogni Stato sta legiferando a passo sostenuto, ma su scala locale e in modo non coordinato, nonostante le caratteristiche delle tecnologie digitali si adattino difficilmente ai confini statali.

Un problema, già presente nello sviluppo e utilizzo di internet, che rischia di essere ancora più evidente nell'era dell'IA e di avere impatti rilevanti sui sistemi economici globali. I possibili impatti, ad esempio, potrebbero riguardare le disparità competitive derivanti dalla mancanza di un *framework* regolatorio globale e la difficoltà delle imprese, soprattutto se operanti in mercati internazionali, a garantire la conformità con le diverse regole proprie di ciascun sistema statale.

V.2 L'AI ACT

V.2.1 STATO DI ATTUAZIONE

L'attuazione del regolamento riflette l'approccio piramidale adottato dalle istituzioni europee nella regolazione dell'IA, con un trattamento differenziato in funzione della tipologia di rischio determinato dai diversi sistemi di IA. In linea generale,

al fine di consentire l'adeguamento delle imprese del settore agli obblighi introdotti dalla disciplina di nuovo conio, l'art. 113 stabilisce che sebbene il regolamento sia entrato in vigore il 2 agosto 2024 (il ventesimo giorno successivo alla relativa pubblicazione nella Gazzetta ufficiale dell'Unione europea avvenuta in data 12 luglio 2024), il medesimo prevede delle date di applicazione differenziate.

Il regime progressivo e differenziato in merito all'applicazione del regolamento emerge alla luce della differente scansione temporale prevista:

- i divieti per le pratiche di IA classificate tra quelle vietate, sono applicati a decorrere dal 2 febbraio 2025, in ragione dell'esigenza di scongiurare l'utilizzo e l'immissione sul mercato di sistemi di IA suscettibili di determinare rischi inaccettabili nei confronti dei diritti fondamentali;
- a partire dal 2 agosto 2025 è prevista l'applicazione delle disposizioni relative ai requisiti dei nuovi modelli di *General Purpose AI* (per quelli già disponibili sul mercato, la data di applicazione è invece agosto 2027);

- è prevista l'applicazione a decorrere dal 2 agosto 2025 del capo III, sez. 4 relativo alle autorità di notifica e agli organismi notificati, incaricati dello svolgimento di un ruolo essenziale nell'applicazione del regolamento, in quanto aventi la funzione di garantire la conformità dei sistemi di Intelligenza Artificiale ai requisiti stabiliti dal medesimo regolamento e alle regole tecniche che dovranno essere successivamente emanate dagli organismi di normazione europei quali il Comitato europeo di normazione (CEN), il Comitato europeo di normazione elettrotecnica (CENELEC) e l'Istituto europeo delle norme di telecomunicazione (ETSI). La medesima data è prevista per l'attuazione del Capo V, relativo ai modelli di IA per finalità generali, per tali intendendosi i modelli che mostrano una generalità significativa e che sono in grado di eseguire un'ampia gamma di compiti indipendentemente dal modo in cui il modello è immesso sul mercato e che può essere integrato in una varietà di sistemi o applicazioni a valle (cfr. art. 3, par. 63);
- il 2 agosto 2025 è la data a partire dalla quale si applicheranno le norme contenute nel Capo VII relativo alla *governance* dell'IA, tanto interna all'UE - attraverso l'istituzione

dell'Ufficio per l'IA, il Consiglio dell'IA (*AI Board*) e il gruppo di esperti scientifici indipendenti - quanto quella predisposta da parte degli Stati membri. La medesima data è altresì prevista per l'applicazione delle norme di cui al Capo XII (con l'eccezione dell'articolo 101), concernente le sanzioni da irrogare nei confronti degli operatori nell'eventualità di una violazione delle norme del regolamento;

- dal 2 agosto 2026, poi, troveranno applicazione i requisiti previsti dalla normativa per i sistemi di IA ad alto rischio;
- a far data dal 2 agosto 2027 si applicheranno l'art. 6, par. 1 relativo alla classificazione dei sistemi di IA come "ad alto rischio" per sistemi integrati in prodotti regolamentati, e i corrispondenti obblighi disciplinati dal Capo III, sez. 2 e 3.

BOX 3. SPAZI DI SPERIMENTAZIONE NORMATIVA - LE POTENZIALITÀ DELLE *REGULATORY SANDBOX*.

L'AI Act europeo prevede una disciplina generale degli «spazi di sperimentazione normativa» per l'intelligenza artificiale. L'articolo 57 del Regolamento individua tali spazi come il contesto in cui garantire un ambiente controllato per promuovere l'innovazione e facilitare: «lo sviluppo, l'addestramento, la sperimentazione e la convalida di sistemi di IA innovativi

per un periodo di tempo limitato prima della loro immissione sul mercato o della loro messa in servizio conformemente a un piano specifico dello spazio di sperimentazione concordato tra i fornitori o i potenziali fornitori e l'autorità competente» (art. 57, par. 5). Gli spazi di sperimentazione possono anche comprendere: «prove in condizioni reali soggette a controllo nei medesimi spazi».

I principali soggetti coinvolti nel funzionamento delle *sandbox* sono i fornitori (compresi quelli potenziali) dei sistemi di IA e l'autorità competente nazionale, che ha la responsabilità di istituire la *sandbox*.

Più nello specifico, l'AI Act prevede che gli Stati membri provvedano affinché le rispettive loro autorità competenti istituiscano almeno uno spazio di sperimentazione normativa per l'IA a livello nazionale. Ciò entro il 2 agosto 2026 (art. 57, par. 1). Inoltre, l'autorità deve fornire orientamenti ai partecipanti riguardo agli obiettivi e requisiti, coordinare e rendicontare l'iniziativa, collaborare con altre autorità nazionali, come quelle per la protezione dei dati personali, e cooperare con le altre autorità a livello europeo nel quadro dell'AI *Board*.

Come per la declinazione di altre misure previste dal regolamento, nella disciplina delle *sandbox* regolatorie l'AI Act ha un'attenzione particolare per le piccole-medie imprese. Viene previsto (all'art. 61, par. 1, lett. a) che le PMI (comprese le start-up), con sede legale o una filiale nell'Unione e che soddisfino le condizioni di ammissibilità e i criteri di selezione, abbiano un accesso prioritario alle *sandbox* per l'IA. Più in generale, devono essere creati canali specifici di comunicazione e consulenza e di agevolazione per le PMI in relazione agli strumenti degli spazi di sperimentazione e al processo di sviluppo della normazione.

Nel corso del 2024, la Commissione europea ha presentato delle iniziative volte allo sviluppo delle *sandbox*, come il bando di finanziamento *DIGITAL-2024-AI-ACT-06-SANDBOX - AI regulatory sandboxes: EU level coordination and support*[270], nell'ambito del programma *Digital Europe*. Anche in questo caso, la proposta della Commissione mostra un'attenzione particolare per le PMI: uno degli obiettivi fondamentali dei progetti finanziati deve essere lo studio della rimozione delle barriere di accesso alle *regulatory sandbox* per PMI e startup, le quali devono essere direttamente coinvolte nei test sviluppati nell'ambito del programma.

Le *sandbox* hanno diverse applicazioni all'interno degli Stati membri anche al di fuori dell'attuazione dell'AI Act. Uno dei primi esempi di *sandbox* risale al 2015 ed è stato istituito in ambito *FinTech* da parte della Financial Conduct Authority (FCA) del Regno Unito[271].

In Italia, vi è una disciplina specifica sulla sperimentazione delle tecnologie *FinTech* con sperimentazione in costante dialogo con Banca d'Italia, CONSOB e IVASS (cfr. art. 36 D.L. 30 aprile 2019, n. 34[272], e il relativo D.M. attuativo 30

[270] Per il bando Digital Euro Program (15 febbraio 2024), cfr. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/digital/wp-call/2024/call-fiche_digital-2024-ai-act-06_en.pdf.

[271] FINANCIAL CONDUCT AUTHORITY, *Regulatory sandbox*, November 2015, https://www.fca.org.uk/publication/research/regulatory-sandbox.pdf. L'apertura delle adesioni è avvenuta a giugno 2016: "Financial Conduct Authority's regulatory sandbox opens to applications", Press Releases FCA, 9 maggio 2016, https://www.fca.org.uk/news/press-releases/financial-conduct-authority%E2%80%99s-regulatory-sandbox-opens-applications.

[272] Consultabile qui: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legge:2019-04-30;34!vig=2025-05-07. Per completezza, si precisa che, al momento in cui si redige il rapporto, si è da poco conclusa (16 maggio 2025) una consultazione sul nuovo schema di regolamento attuativo dell'art. 36 citato del D.L. 34/2019, sul funzionamento del Comitato e della sperimentazione *FinTech*, predisposto con la finalità principale di semplificare il procedimento di ammissione alla *sandbox*.
Per la pagina relativa alla consultazione con tutti i documenti relativi cfr. https://www.dt.mef.gov.it/it/dipartimento/consultazioni_pubbliche/consultazioni_in_corso/consultazione_fintech/index.html.

aprile 2021, n. 100[273]), in esecuzione della quale è stata aperta la "seconda finestra di sperimentazione a fine 2023[274]. Più in generale, l'art. 36, D.L. n. 76/2020 (cd. Decreto semplificazioni 2020[275]), prevede la possibilità di una deroga temporanea a norme che impediscano la «sperimentazione di idee e iniziative» innovative in ambito tecnologico e di digitalizzazione: «volte al miglioramento della competitività, dell'efficienza e dell'efficacia di servizi a cui cittadini e imprese».

Infine, è da segnalare una proposta di legge presentata dal governo tedesco[276] nel novembre 2024 dedicata specificamente al miglioramento delle condizioni quadro per la sperimentazione delle innovazioni nelle *sandbox* regolatorie e la promozione dell'apprendimento normativo (cd. *Reallabore-Gesetz*).

V.2.2 LE INIZIATIVE DELLA COMMISSIONE: *CODE OF PRACTICE* SU *AI GENERAL-PURPOSE*, DEFINIZIONE SISTEMA DI IA E LINEE GUIDA SU PRATICHE PROIBITE

In considerazione dell'impatto che possono avere i modelli di IA per finalità generali (GPAI o *General Purpose Artificial Intelligence*)

[273] Il D.M. è consultabile qui: https://www.normattiva.it/uri-res/N2Ls?urn:nir:ministero.economia.e.finanze:decreto:2021-04-30;100!vig=2025-05-07.

[274] La pagina esplicativa di Banca d'Italia sulle modalità di accesso e la seconda fase di sperimentazione è reperibile qui: https://www.bancaditalia.it/focus/sandbox/ammissione-sperimentazione/index.html?dotcache=refresh

[275] Norma già citata, v. nota 272.

[276] Per la proposta, cfr. https://www.bmwk.de/Redaktion/DE/Downloads/Gesetz/20241113-breg-reallaboreg-download.pdf?__blob=publicationFile&v=7. Una sintesi in inglese della proposta e di altre iniziative del Legislatore e del Governo federale tedesco è reperibile qui: https://www.bmwk.de/Redaktion/EN/Dossier/regulatory-sandboxes.html.

la Commissione europea, per il tramite dell'Ufficio europeo per l'IA, sta procedendo all'elaborazione di un codice di condotta (*code of practice*) in ragione dei rischi sistemici nei quali si incorre in tale ambito.

Ai sensi dell'art. 3, par. 63 dell'AI Act, questi ultimi sono quei modelli che mostrano una generalità significativa e che sono in grado di eseguire un'ampia gamma di compiti, indipendentemente dal modo in cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle.

Al fine di garantire che l'IA sia affidabile, sicura e rispettosa dei diritti umani, l'AI Act ha stabilito una serie di requisiti per i fornitori di tali modelli, quali obblighi di trasparenza e di tutela del diritto d'autore. Il codice di condotta persegue l'obiettivo di chiarire e specificare tali requisiti, nel tentativo di agevolare anche l'attività di conformità normativa da parte dei relativi fornitori.

Il 30 settembre 2024, l'AI Office ha ospitato il "*Kick-off Event*" al quale hanno preso parte mille soggetti tra esperti e organizzazioni[277].

[277] Si vedano gli aggiornamenti alla pagina della CE, General-Purpose AI Code of Practice: https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice.

Quattro gruppi di lavoro, ciascuno competente per specifici ambiti, compongono la struttura della plenaria; le relative attività - la prima delle quali si è svolta il 23 ottobre 2024[278] - sono improntate alla massima trasparenza e inclusività, dal momento che sono coinvolti vari portatori di interessi come: i fornitori di tali modelli, esponenti della società civile, accademici ed esperti indipendenti. Il ruolo dell'AI Office è quello di supervisionare, coordinare e coadiuvare l'attività dei gruppi di lavoro.

Il 14 novembre 2024, l'AI Office ha pubblicato la prima bozza del codice di condotta[279], nella quale sono indicati gli obiettivi da raggiungere, le misure utili a tal fine e i c.dd. "*key performance indicators*" (KPI). L'obiettivo principale è quello di consegnare ai fornitori dei sistemi di IA per finalità generali un *vademecum* utile per conformarsi ai requisiti prescritti dall'AI Act, facilitando al contempo l'attività di controllo da parte dell'Ufficio per l'IA.

[278] General-Purpose AI Code of Practice, *ibidem*.

[279] EUROPEAN COMMISSION, *First Draft of the General-Purpose AI Code of Practice*, written by independent experts, 14 November 2024, https://digital-strategy.ec.europa.eu/en/library/first-draft-general-purpose-ai-code-practice-published-written-independent-experts. A tal proposito cfr. "US tech giants ask European Commission for 'simplest possible' AI code", *Euronews*, 30 May 2025, https://www.euronews.com/next/2025/05/30/us-tech-giants-ask-european-commission-for-simplest-possible-ai-code.

Secondo la *timeline* indicata dalla Commissione, la versione finale del codice di condotta dovrebbe essere approvata in un'assemblea plenaria degli *stakeholders* dedicati e pubblicata entro agosto 2025, per poi essere adottata dalla Commissione tramite un atto di esecuzione. Al momento la Commissione si sta confrontando con gli *stakeholders* (incluse imprese *Big Tech*) sulla terza bozza del testo presentata a marzo 2025.

L'attività regolatoria della Commissione è particolarmente rilevante, essendo quest'ultima titolare del potere di adottare *guidelines* (orientamenti) non vincolanti ai fini della corretta e uniforme interpretazione delle disposizioni dell'AI Act.

Nel dettaglio, al fine di fornire un quadro completo in tema di sistemi di intelligenza artificiale per scopi generali (la già ricordata GPAI), la Commissione europea pubblicherà alcuni orientamenti a seguire la consultazione aperta, dove le risposte degli *stakeholders* son state possibili sino al 22 maggio 2025[280]. Mentre anche la pubblicazione degli orientamenti è prevista per agosto 2025.

[280] Cfr. "Commission seeks input to clarify rules for general-purpose AI models", Press release, 22 April 2025, https://digital-strategy.ec.europa.eu/en/news/commission-seeks-input-clarify-rules-general-purpose-ai-models.

Inoltre, si segnala che in data 4 febbraio 2025 - e nell'esercizio dei poteri espressamente conferitele dall'art. 96, par. 1, lett. b dell'AI Act - la Commissione europea ha adottato le *guidelines* (orientamenti) in merito alle pratiche proibite di intelligenza artificiale[281], stabilite dall'art. 5 dell'AI Act. Ciò al fine di soddisfare esigenze di certezza del diritto attraverso l'uniforme interpretazione delle disposizioni del regolamento, agevolando al contempo l'attività di conformità normativa da parte delle imprese del settore.

Fatta la premessa che l'accertamento di un sistema vietato costituisce l'esito di una valutazione casistica da compiersi alla luce delle circostanze del caso concreto, la Commissione ha indicato una lista di pratiche vietate, in via esemplificativa e non esaustiva. Così rientrano nel divieto di cui all'art. 5 i sistemi manipolativi, di controllo sociale e di sorveglianza di massa, in quanto lesivi dei diritti fondamentali protetti dalla Carta di Nizza, quali il diritto a non essere discriminati, all'eguaglianza,

[281] "La Commissione pubblica gli orientamenti sulle pratiche vietate in materia di intelligenza artificiale (IA), quali definite dalla legge sull'IA", 4 febbraio 2025, https://digital-strategy.ec.europa.eu/it/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act.

alla protezione dei dati personali, il rispetto della vita privata e familiare, la libertà di espressione e informazione, la libertà di assemblea. Tra queste pratiche sono individuati i sistemi di *social scoring*, la raccolta indiscriminata di dati biometrici dal *web* per l'identificazione facciale e i sistemi di rilevazione delle emozioni.

L'ambito applicativo del divieto è circoscritto alla immissione sul mercato, la messa in servizio e all'uso di tali sistemi: l'immissione sul mercato consiste nella prima messa a disposizione del sistema nel mercato UE, per la distribuzione e l'uso nell'ambito di un'attività commerciale, indipendentemente dal fatto che quest'ultima sia svolta a titolo oneroso o gratuito.

La messa in servizio consiste nella fornitura per la prima volta in favore del deployer o per un utilizzo proprio nel territorio dell'Unione. L'utilizzo va inteso in maniera ampia, come l'impiego del sistema di IA in qualsiasi fase del ciclo di vita del medesimo, in seguito alla propria immissione sul mercato o alla messa in servizio.

Sebbene i soggetti destinatari delle disposizioni dell'AI Act siano eterogenei, la Commissione ha avuto premura di specificare che i divieti di cui all'art. 5 si applicano nei soli confronti dei fornitori

e dei *deployers*, e non anche per gli importatori, i distributori e i produttori. Le responsabilità tra il fornitore e il *deployer* sono distribuite in base ai propri ruoli nella catena del valore dei sistemi di IA e sulla scorta dei rispettivi poteri di controllo.

Ai fini della perimetrazione dell'ambito applicativo del regolamento, sono chiariti i confini tra le attività funzionali a perseguire obiettivi di sicurezza o di difesa nazionale e militari da un lato e le iniziative di contrasto dall'altro: applicando le medesime coordinate tracciate dalla Corte di giustizia dell'Unione europea[282], sono meglio indicate le caratteristiche della prima tipologia di attività, le quali soltanto sono escluse dall'applicazione dell'AI Act.

La Commissione coglie l'occasione di chiarire il regime normativo applicabile nell'intervallo temporale che intercorre tra il momento in cui i divieti di cui all'art. 5 trovano applicazione (2 febbraio 2025) e quello in cui produrranno i propri effetti le disposizioni in materia di controllo da parte delle autorità

[282] In particolare, nelle pronunce della Corte di Giustizia UE 6 ottobre 2020, *La Quadrature du Net*, C-511/18, C-512/18 e C-520/18; 5 giugno 2023, *Commissione c. Polonia*, C-204/21; 22 giugno 2021, *Latvijas Republikas Saeima*, C-439/19; 6 ottobre 2020, *Privacy International*, C-623/17.

nazionali ed europee e quelle concernenti le sanzioni in caso di inosservanza del regolamento, il 2 agosto 2025.

A tal riguardo, è specificato che i divieti sono pienamente operativi e vincolanti sia per i fornitori che per i deployer, i quali saranno dunque tenuti ad assicurare che non siano immessi sul mercato, né utilizzati i sistemi riconducibili alla lista di cui all'art. 5. Ciò comporta che coloro che si ritengono pregiudicati da tali pratiche prima del 2 agosto 2025 non sono completamente sforniti di tutela, bensì potranno adire le autorità nazionali, tanto amministrative quanto giurisdizionali, ai fini dell'ottenimento di misure inibitorie contro siffatte pratiche.

In data 6 febbraio 2025 - nell'esercizio dei poteri di cui all'art. 96, par. 1, lett. f dell'AI Act - la Commissione ha adottato le proprie *guidelines* (orientamenti) concernenti la definizione dei sistemi di IA[283] di cui all'art. 3, par. 1 del medesimo regolamento,

[283] "La Commissione pubblica orientamenti sulla definizione del sistema di IA per facilitare l'applicazione delle norme della prima legge sull'IA", 6 febbraio 2025, https://digital-strategy.ec.europa.eu/it/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application.

contribuendo a chiarirne il relativo perimetro applicativo. La premessa da cui muove la Commissione risiede nella consapevolezza che non è possibile dettare una classificazione esaustiva dei sistemi di IA, come chiarito dallo stesso regolamento, il quale poneva già l'accento sulla necessità di un approccio elastico in grado di fronteggiare i rapidi sviluppi tecnologici che permeano il settore.

Sono così enfatizzati - in conformità al lavoro già svolto dall'Organizzazione per la Cooperazione e lo Sviluppo Economico nel proprio *Explanatory memorandum* del 5 marzo 2024[284] - i sette requisiti che devono o possono caratterizzare il funzionamento di un sistema di IA per essere definito tale, ovvero: 1.) deve trattarsi di un sistema automatizzato e dunque fondato su un meccanismo computazionale; 2.) deve essere designato per operare con livelli di autonomia variabile, in quanto tale in maniera indipendente rispetto all'operatore umano; 3.) può presentare adattabilità dopo la relativa

[284] Cfr. OECD (2024), "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en.

diffusione, consentendo al sistema di cambiare il proprio funzionamento producendo sempre nuovi output pur partendo dai medesimi input; 4.) deve funzionare per obiettivi espliciti o impliciti, a seconda che questi ultimi siano impartiti dallo sviluppatore o siano desumibili dalle istruzioni fornite al sistema e dall'interazione con l'ambiente esterno; 5.) deve possedere capacità inferenziale, essendo capace di generare contenuti di varia natura in funzione dei dati impartiti dallo sviluppatore attraverso tecniche peculiari, individuate dalla medesima Commissione in via esemplificativa (*machine learning*, apprendimento supervisionato e non, *deep learning*); 6.) gli output prodotti possono essere di varia natura e consistere in previsioni, contenuti, raccomandazioni o decisioni e 7.) capaci di influenzare ambienti fisici o virtuali.

BOX 4. IL REGOLAMENTO EUROPEO IN MATERIA DI INTELLIGENZA ARTIFICIALE E IL PROBLEMA DEFINITORIO.

La pubblicazione del regolamento europeo in materia di intelligenza artificiale (qui nel *box* in sigla, RIA, altresì noto come AI Act), sebbene sia motivo di orgoglio e autogratificazione delle Istituzioni europee, è terreno fertile per serie e critiche riflessioni.

In primo luogo, optando per una legislazione di tipo orizzontale, il legislatore europeo dichiara di aver accolto l'idea alla base della neutralità tecnologica ovvero che la tecnologia debba essere regolamentata, indipendentemente dalle proprie specificità. L'espressione assai ricorrente nel discorso di politica del diritto è, tuttavia, foriera di incertezze, potendo assumere significati e connotazioni differenti. Da un lato, infatti, non si vuole essere neutrali rispetto alle tecnologie poiché non tutte sono egualmente commendevoli. Da un altro lato, invece, si vorrebbe perseguire tale neutralità per assicurare una maggiore elasticità del sistema, rendendolo cioè *future proof*.

L'idea sottesa è quella per cui lo sviluppo del quadro normativo non potrebbe mantenere il passo del ben più rapido sviluppo tecnologico. La neutralità così intesa servirebbe allora il fine di rendere le norme adattabili in quanto non *technology specific*.

Anche siffatta petizione di principio che, ad una lettura piana, potrebbe apparire condivisibile è, invece, assai problematica nella misura in cui forza una semplificazione e aggregazione di applicazioni tra loro diversissime, senza offrire un chiaro criterio discretivo.

L'intelligenza artificiale non esiste: non esiste cioè un solo tipo di applicazioni, una caratteristica davvero essenziale e ricorrente, capace di costituire un chiaro filo rosso che tenga assieme tra loro applicazioni diversissime, come un algoritmo per la negoziazione di strumenti finanziari ad alta frequenza e uno spazzolino elettrico capace di verificare il corretto utilizzo dello strumento da parte dell'utente.
Il termine IA è, nei fatti, un termine ombrello, usato per richiamare in modo sintetico e mai davvero preciso e definito applicazioni che presentano criteri del tutto opposti, autonome - come un veicolo o un drone - oppure no - come un robot chirurgico, una protesi robotica o un esoscheletro, intelligenti - come

un sistema di diagnostica medica - oppure no - come un aspirapolvere automatico o uno spazzolino, capaci di imparare - come un *large language model* - oppure no - come buona parte dei robot industriali.
A ben vedere poi, anche là dove due o più applicazioni possono presentare un tratto comune (p.e. la capacità di apprendere) esso può declinarsi in modi tanto diversi da non poter davvero giustificare l'accostamento, se non tralasciando aspetti dirimenti rispetto tanto alle soluzioni tecniche messe in opera quanto alle capacità effettivamente conferite.

In questo senso, una eccessiva livellazione ai fini regolatori potrebbe rischiare di costituire una semplificazione eccessiva, incapace di cogliere quelle differenze che suscitano poi la necessità di adottare soluzioni regolatorie ancora fortemente distinte se non del tutto opposte.

L'impostazione di un approccio regolatorio orizzontale rischia di trovare un ostacolo nella complessità del reale ed esige - pena l'assoluta inadeguatezza rispetto al proprio fine di governare l'innovazione - un contemperamento costituito dall'introduzione di un criterio discretivo ulteriore. Il regolamento fa dunque ricorso ad una classificazione per livelli di rischio, inteso come danno ai diritti fondamentali, alla salute e alla sicurezza. Tuttavia, tale classificazione, in assenza di una 'clusterizzazione' di tipo verticale potrebbe non essere né sufficiente, né di grande aiuto.

L'AI Act suddivide i sistemi di IA in base a diversi livelli di rischio che i sistemi medesimi possono generare: i sistemi di IA a rischio inaccettabile (proibiti), ad alto rischio (consentiti, ma oggetto di determinati obblighi e requisiti) e, per differenza, quelli non ad alto rischio (compresi quelli considerati a rischio minimo o nullo - consentiti, senza restrizioni). A seconda

del livello di rischio sono predisposte a carico del produttore/sviluppatore/*deployer* una serie di obbligazioni più o meno stringenti.

Ora, rispetto alla nozione di rischio, preme tuttavia sottolineare come essa non sia né definita, né misurata e non è neppure fornito un criterio o un metodo che giustifichi per quale motivo proprio le applicazioni di cui all'art. 5 RIA dovrebbero comportare un rischio del tutto inaccettabile e non altre.
L'impostazione adottata è "*risk-based*" e incoraggia un approccio che si fonda sulla tutela di libertà e diritti fondamentali alla base del patrimonio culturale della stessa Europa.

Si badi che ciò non equivale a criticare l'esistenza stessa di una classificazione né, tantomeno, la necessità di proibire alcuni usi particolarmente pericolosi di queste tecnologie. Piuttosto, la stessa formulazione dell'art. 5 RIA potrebbe apparire insufficiente a prevenire alcune forme di lesioni significative di interessi giuridicamente rilevanti.

Tuttavia, il legislatore europeo non individua criteri obiettivi e neppure opera una classificazione degli interessi della persona la cui lesione ritiene particolarmente significativa, più significativa di altre.

In questo senso, ritenere che le applicazioni di cui all'art. 5 RIA presentino un rischio necessariamente maggiore rispetto alle applicazioni di cui all'art. 6 RIA (le applicazioni c.d. ad alto rischio) operata dal legislatore è certamente ammissibile, eppure non sembra avere a che vedere con una effettiva misura del diverso livello di rischio.
Lo stesso richiamo al rischio può apparire allora un elemento fuorviante e non certo il vero criterio sulla scorta del quale - attraverso un ragionamento oggettivo, ben definito e necessitato nelle proprie conclusioni - si arrivi a concludere che quelle e solo quelle applicazioni meritano di essere qualificate

come proibite, mentre altre sono semplicemente sufficientemente rischiose da meritare una disciplina di particolare rigore.

Ciò premesso, se ci si addentra nelle definizioni di cui all'art. 5 e 6 RIA, che certamente costituiscono il cuore dell'intera disciplina, si può evidenziare come esse non offrano la completa certezza tanto all'interprete giurista quanto al tecnico sviluppatore o utilizzatore di una applicazione di IA.

Per quanto riguarda l'art. 5 RIA, chiamato a delineare le c.d. pratiche proibite, la Commissione Europea ha ritenuto di dover pubblicare delle *guidelines* (orientamenti) di quasi 150 pagine (di cui si è detto sopra, nel par. V.2.2), per meglio chiarire tali definizioni ambigue. Se, da un lato, si può dubitare del valore di tali indicazioni su di un piano applicativo - nella misura in cui un giudice non sarebbe tenuto a rispettarne il contenuto in quanto certamente non cogente - da un altro lato, anche una lettura attenta delle medesime potrebbe non essere in grado di fugare talune incertezze circa le applicazioni che, anche a posteriori[285] e dopo la loro immissione sul mercato, potrebbero essere considerate proibite. L'incertezza per potenziali sviluppatori è dunque significativa e certamente tale da scoraggiare potenzialmente l'innovazione.

Per quanto attiene l'art. 6 RIA, chiamato a definire le applicazioni ad alto rischio e soggette, dunque, a vincoli regolatori maggiori, anch'esso lascia residuare taluni profili di incertezza.

[285] In particolare, l'art. 5, §1, let. (a) e (b) per definire quali usi siano proibiti fa riferimento alla circostanza che le applicazioni manipolatorie o tali da sfruttare le particolari vulnerabilità individuali siano tali da cagionare un danno significativo al singolo o a gruppi di soggetti. Ora, perché si possa apprezzare un danno significativo è necessario che l'applicazione sia commercializzata e utilizzata per un periodo di tempo sufficiente, dovendosi con ciò concludere che un sistema immesso sul mercato nella convinzione della sua piena liceità potrebbe essere, a posteriori - quando i costi di ricerca, sviluppo, produzione e commercializzazione sono stati sostenuti - considerato illecito e ritirato dal mercato.

In particolare, da un lato, prevede il richiamo ad alcuni settori di applicazione (art. 6, §2 e Allegato III RIA) che sono definiti in modo molto ampio, prevedendo altresì un complesso intreccio di regole ed eccezioni che rende complesso tracciare una linea netta tra sistemi potenzialmente assai simili. Da un altro lato, richiama una buona parte delle discipline in materia di sicurezza dei prodotti (art. 6, §1 e Allegato I RIA), stabilendo che il sistema sarà ad alto rischio laddove esso stesso ovvero una sua componente di sicurezza sia dotata di IA e sia soggetta alla certificazione mediante intervento di un ente terzo notificato.

Ora, mancando una definizione di componente di sicurezza, sembra difficile determinare quale componente sia in grado di giustificare la classificazione del sistema come ad alto rischio. Non è chiaro, ad esempio, se il sistema di identificazione degli oggetti di un robot mobile, che faccia ricorso ad un qualche algoritmo, debba essere considerato un sistema di sicurezza e non invece il solo impianto frenante o sterzante.

Il quadro qui sinteticamente tratteggiato fa emergere incertezze definitorie capaci di produrre potenziali conseguenze sul piano applicativo e qualificatorio delle tecnologie.

È dunque ragionevole presupporre un elevato grado di incertezza normativa *ex ante*, e di potenziale contenzioso *ex post*, rispetto al quale solo soggetti dotati di rilevante potere economico e di mercato (cd. *big tech*) potranno rimanere sostanzialmente indifferenti, eventualmente consolidando la propria posizione di dominanza. In questo senso, la stessa capacità di governo dell'innovazione potrebbe risultare profondamente compromessa.

BOX 5. LA CONSULTAZIONE MIRATA DELLA DG FISMA SULL'USO DELL'IA NEI SERVIZI FINANZIARI.

Nel giugno del 2024, la Commissione europea (DG FISMA) ha lanciato una consultazione mirata sull'uso dell'intelligenza artificiale (IA) nei servizi finanziari[286], con l'obiettivo di raccogliere opinioni da parte degli stakeholder del settore. La consultazione aveva come obiettivo l'individuazione dei principali casi d'uso e di benefici, barriere e rischi connessi all'adozione dell'IA in ambito finanziario, fornendo così un sostegno per la Commissione nella valutazione dello sviluppo e dei rischi del mercato in relazione all'IA e per l'attuazione efficace dell'AI Act nel settore finanziario, oltre che dei framework legislativi a esso attinenti.

Nel documento di avvio della consultazione, la Commissione osserva come nel settore finanziario, l'AI Act si debba integrare con il quadro normativo esistente, che include la Direttiva MiFID (2014/65/EU), la Direttiva sul credito ai consumatori, le normative antiriciclaggio (AMLD) e altre disposizioni sui mercati finanziari.

Tra gli aspetti principali relativi al settore finanziario, l'AI Act identifica due aree di applicazione ad alto rischio (all'allegato II): (i.) sistemi di IA per la valutazione del merito creditizio e assegnazione di punteggi di credito ai consumatori, con l'eccezione di quelli usati esclusivamente per rilevare frodi

[286] Cfr. COMMISSIONE EUROPEA, DIREZIONE GENERALE DELLA STABILITÀ FINANZIARIA, DEI SERVIZI FINANZIARI E DELL'UNIONE DEI MERCATI DEI CAPITALI, *Targeted Consultation on Artificial Intelligence in the Financial Sector*, Reference Documents, 2024, https://finance.ec.europa.eu/regulation-and-supervision/consultations-0/targeted-consultation-artificial-intelligence-financial-sector_en; Per l'Italia, si rimanda alla Unità di informazione finanziaria di Banca d'Italia e i suoi *Quaderni dell'antiriciclaggio* e alla fonte riportata sul sito del MEF riguardo l'uso dell'IA nelle attività antiriciclaggio: https://www.dt.mef.gov.it/export/sites/sitodt/modules/documenti_it/news/news/Data_science_and_social_network_analysis_for_Anti-Money_Laundering.pdf.

finanziarie e (ii.) sistemi di IA per la valutazione del rischio e la determinazione dei prezzi nell'assicurazione vita e sanitaria, che possono influenzare l'accesso alle coperture assicurative per le persone fisiche.

La consultazione è strutturata in tre sezioni principali:

1. domande generali sull'IA nei servizi finanziari: questa parte esplora i benefici dell'IA (come rilevamento delle frodi, automazione, riduzione dei costi e personalizzazione dei servizi) e le sfide da affrontare (rischi di *bias*, trasparenza degli algoritmi, compliance normativa e sicurezza informatica).
2. Casi d'uso specifici dell'IA nel settore finanziario: sono qui analizzati gli impatti dell'IA nei seguenti ambiti: banche e pagamenti, *scoring* del credito, gestione del rischio, *compliance* normativa, prevenzione delle frodi, assistenza clienti; mercati finanziari e infrastrutture di mercato: *trading* algoritmico, rilevazione di abusi di mercato, gestione della liquidità; assicurazioni e pensioni: *pricing* e sottoscrizione delle polizze, gestione dei sinistri, rilevazione delle frodi; gestione patrimoniale e *asset management*; gestione del rischio, *robo-advisor*, ottimizzazione dei portafogli di investimento.
3. Applicazione dell'AI Act nel settore finanziario: Questa sezione raccoglie input sulle esigenze del settore per la corretta attuazione delle norme dell'AI Act, in particolare per le applicazioni ad alto rischio. Gli operatori del settore sono invitati a esprimere il loro punto di vista su eventuali sfide normative e sulle esigenze di chiarimento delle disposizioni di legge.

La consultazione si è chiusa il 13 settembre 2024 e i suoi risultati non sono ancora stati resi pubblici.

V.3 IL DATA ACT E LE FAQ DELLA COMMISSIONE

Le previsioni del Data Act (Regolamento (UE) 2023/2854[287], pubblicato in Gazzetta Ufficiale UE il 22 dicembre 2023 ed entrato in vigore l'11 gennaio 2024) saranno applicabili dal 12 settembre 2025[288]. Il Data Act si pone come tassello fondamentale della regolazione del mercato unico europeo per la libera circolazione dei dati, insieme ad altre iniziative quali il Data Governance Act (Regolamento (UE) 2022/868).

Con questo obiettivo di fondo, il Data Act regola, tra l'altro, la condivisione dei dati generati dall'uso di prodotti connessi e i servizi correlati (nel contesto quindi dell'internet delle cose, cd.

[287] Per esteso: Regolamento (UE) 2023/2854 del Parlamento europeo e del Consiglio, del 13 dicembre 2023, riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo e che modifica il regolamento (UE) 2017/2394 e la direttiva (UE) 2020/1828 (regolamento sui dati), https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=OJ%3AL_202302854.

[288] Alcune previsioni del Regolamento si applicano a partire da altre scadenze (cfr. art. 50: "L'obbligo derivante dall'articolo 3, paragrafo 1, si applica ai prodotti connessi e ai servizi correlati immessi sul mercato dopo il 12 settembre 2026.

Il capo III si applica solo in relazione agli obblighi di messa a disposizione dei dati a norma del diritto dell'Unione o della legislazione nazionale adottata in conformità del diritto dell'Unione, che entrano in vigore dopo il 12 settembre 2025.

Il capo IV si applica ai contratti conclusi dopo il 12 settembre 2025.

Il capo V si applica a decorrere dal 12 settembre 2027 ai contratti conclusi il o anteriormente al 12 settembre 2025, a condizione che:

a.) siano a tempo indeterminato; o

b.) scadano almeno dieci anni dopo l'11 gennaio 2024").

IOT), segnatamente a favore degli utenti di tali servizi e delle imprese, con un'attenzione particolare per le PMI.

L'iniziativa è di rilevanza anche per lo sviluppo di sistemi di intelligenza artificiale nel mercato europeo, in quanto permette l'accesso a dati con standard di condivisione predefiniti: i dati sono, come ribadito anche nel contesto dell'AI Act, una delle due componenti fondamentali - insieme agli algoritmi - per lo sviluppo di sistemi di IA.

Nel corso della seconda metà del 2024, la Commissione europea ha pubblicato una lista di risposte a "domande frequenti" (FAQ)[289], le quali offrono diverse chiavi interpretative rispetto al contenuto delle norme del Data Act. Nell'ultima versione del documento (del 3 febbraio 2025), la Commissione offre indicazioni operative su diversi punti del regolamento.

Anzitutto, le FAQ trattano del rapporto tra il Data Act e le altre normative dell'Unione europea, con un'attenzione particolare per il Regolamento generale sulla protezione dei dati personali

[289] "La Commissione pubblica le domande frequenti sulla legge sui dati", 6 settembre 2024, https://digital-strategy.ec.europa.eu/it/library/commission-publishes-frequently-asked-questions-about-data-act.

(Regolamento (UE) 2016/679 - GDPR). La Commissione precisa che il Data Act stabilisce un quadro orizzontale in materia di condivisione di dati, complementare ai regolamenti settoriali già esistenti. Con riguardo al GDPR, il Data Act ne integra il contenuto in relazione alla gestione dei dati personali, ma non ne modifica le disposizioni. In caso di conflitto tra i due regolamenti, è il GDPR a prevalere, come previsto da una disposizione dello stesso Data Act (art. 1 par. 5).

Diverse altre FAQ riguardano l'ambito di applicazione oggettivo e soggettivo delle norme del Data Act. Si pone, ad esempio, il problema di identificare quali dati in concreto siano da considerare "dedotti" o "ricavati", qualifica che determina la mancata applicazione degli obblighi del Data Act.

A tal riguardo, le FAQ chiariscono che gli elementi per determinare la distinzione tra dati in forma "primaria" (soggetti alle regole del Data Act) o dati "arricchiti" (non soggetti alle regole) si riscontrano nel Considerando 15 del Regolamento: in questo senso, la presenza di "modifiche sostanziali", o di "investimenti sostanziali" (che coinvolgono anche l'uso di "algoritmi proprietari complessi") può determinare l'esclusione

dagli obblighi di condivisione. Nello stesso senso, le FAQ offrono indicazioni anche sul significato di "contenuti" (anch'essi esclusi dall'applicazione del Data Act), "prodotti connessi" e "servizi correlati" (questi ultimi sono termini centrali per determinare l'ambito applicativo del Regolamento).

Per quanto riguarda l'ambito di applicazione soggettivo, le FAQ chiariscono alcuni aspetti relativi alla definizione di "utente" (quali il caso di utenti fuori dai confini dell'UE e di utenti multipli), "titolare dei dati" (ad esempio, quando un fabbricante di prodotti connessi può non essere considerato tale) e terze parti "destinatari dei dati" (precisando, tra l'altro, l'esclusione dei *gatekeeper* - ai sensi del Regolamento (UE) 2022/1925 *DMA* - da tale categoria di soggetti in forza delle regole del Data Act).

Altre indicazioni rilevanti riguardano le modalità di esercizio dei diritti da parte degli utenti: i quali possono chiedere ai titolari di trasferire i dati stessi a terze parti designate (art. 5). Sono anche chiarite le modalità attraverso cui i titolari possono chiedere un compenso "non discriminatorio e ragionevole" ai terzi destinatari di tali dati per la messa a disposizione dei

medesimi per servizi a valore aggiunto in una relazione "*business to business*" (art. 9). Con riguardo a quest'ultimo punto, le FAQ chiariscono che non vi sono importi massimi o minimi di riferimento, ma il calcolo per la determinazione del compenso deve essere eseguito secondo criteri oggettivi (p. es. costi sostenuti e quantità di dati messi a disposizione) e descritti in modo trasparente, mentre non può includere un profitto nel caso in cui i destinatari siano PMI o organizzazioni di ricerca *no profit*.

Altri elementi della disciplina chiariti dalla Commissione riguardano i limiti entro cui la condivisione può essere negata (ossia in caso di seri rischi per la sicurezza o danni economici gravi), gli altri obblighi di condizioni "eque, ragionevoli e non discriminatorie" ("FRAND") per la messa a disposizione dei dati (oltre al compenso), i rimedi in caso di controversie tra i soggetti coinvolti - inclusa la possibilità di rivolgersi ad un organismo per la risoluzione delle controversie specificamente individuato dallo Stato Membro di riferimento - il contenuto delle clausole contrattuali abusive, nei contratti relativi alla condivisione dei dati.

Le FAQ si dedicano estesamente anche alla disciplina relativa alla richiesta da parte di enti pubblici di messa a disposizione dei dati sulla base di necessità eccezionali (capo V del Data Act), chiarendo le condizioni e i limiti entro cui tali dati possono essere richiesti e condivisi con altri enti (art. 21 - per attività di analisi o ricerca scientifica compatibile con la finalità della richiesta, nonché per la produzione di statistiche da parte di Istituti nazionali di statistica o Eurostat).

Un ulteriore capo cui le FAQ dedicano specifici chiarimenti è quello relativo al passaggio tra servizi di trattamento dei dati (capo VI). In questo contesto, la Commissione chiarisce la differenza tra i termini "dati esportabili" e "risorse digitali", entrambi già oggetto di definizione nel Data Act, evidenziando come la seconda includa anche applicazioni e tecnologie per usare effettivamente i dati trasferiti. Sono poi chiarite le modalità di trasferimento dei dati (anche attraverso gli standard tecnici di interoperabilità che saranno messi a disposizione dalla Commissione nell'"Archivio centrale dell'Unione" e le clausole

contrattuali standard relativi al *cloud computing*) e l'applicazione del divieto (dal 12 gennaio 2027) di imporre tariffe di passaggio.

La Commissione chiarisce, inoltre, i contenuti delle misure per impedire l'accesso illegittimo a dati non personali da parte di autorità extra UE e la disciplina del trasferimento internazionale dei dati (in relazione al Capo VII del Data Act) sull'interoperabilità (Capo VIII) e sull'attuazione tramite specifiche autorità individuate dagli Stati membri (Capo IX).

Infine, la Commissione intende offrire indicazioni su alcuni strumenti previsti dal Data Act per dare informazioni utili alla sua applicazione. Nella specie si tratta di:

- linee guida sul compenso ragionevole per la messa a disposizione dei dati, su cui dovrà esprimersi il Comitato per l'innovazione digitale (EDIB) previsto dallo stesso regolamento, una volta che quest'ultimo diverrà applicabile (quindi, come visto, dal 12 settembre 2025),

· richiesta di standardizzazione nominata "*European Trusted Data Framework*", volta alla produzione di standard utili all'applicazione del Data Act[290];

· clausole contrattuali tipo relative a *data sharing* e al loro relativo utilizzo e clausole contrattuali standard per i contratti di *cloud computing* (entrambe previste dall'articolo 41 del Data Act come riferimento non vincolante per gli operatori), elaborate da parte dell'*Expert Group on B2B data sharing and cloud computing contracts* (sotto la gestione della DG JUST e della DG CNECT della Commissione) – la bozza delle clausole è disponibile nel *report* finale dell'*Expert Group* del 2 aprile 2025[291].

[290] La richiesta è stata poi inserita come una delle azioni (n. 8) del Programma di lavoro annuale dell'Unione per la normazione europea per il 2025. Cfr. Comunicazione della Commissione, Programma di lavoro annuale dell'Unione per la normazione europea per il 2025 del 27.3.2025, C/2025/1818, https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=OJ:C_202501818.

[291] *Final Report of the Expert Group on B2B data sharing and cloud computing contracts*, 2 April 2025, https://www.aigl.blog/content/files/2025/04/Final-Report-of-the-Expert-Group-on-B2B-data-sharing-and-cloud-computing-contracts.pdf, per il gruppo di esperti v. anche: https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&groupID=3840.

BOX 6. CONDIVISIONE E ACCESSO A DATI FINANZIARI - LA PROPOSTA DI REGOLAMENTO FIDA.

Il procedimento di approvazione della proposta di Regolamento (UE) che crea un quadro per l'accesso ai dati finanziari (cd. "*Financial Data Access"*, in sigla FIDA)[292], presentata dalla Commissione il 28 giugno 2023 (COM/2023/360)) è - al momento in cui si scrive - ancora in corso. La proposta di regolamento mira a introdurre alcune norme per consentire a consumatori e imprese di controllare l'accesso ai propri dati finanziari, permettendo agli utenti (*data users*), previa autorizzazione, e secondo le disposizioni di dettaglio di FIDA, di accedere a tali dati. In quest'ottica, la proposta si pone espressamente come integrazione e specificazione delle regole "orizzontali" del Data Act nel contesto finanziario. Difatti, la proposta FIDA introduce regole armonizzate sull'individuazione dei dati da condividere, le modalità di tale condivisione, il compenso ai titolari dei dati per la loro messa a disposizione, la promozione della trasparenza e della comparabilità tra i dati, il controllo dei clienti sui propri dati e le relative misure di attuazione.

La proposta del Regolamento FIDA è collegata, altresì, alla proposta di regolamento sui servizi di pagamento (PSR), presentata dalla Commissione sempre il 28 giugno 2023 (COM (2023) 367) e che revisiona l'attuale PSD2 (che, a sua volta, già consente l'accesso ai dati di pagamento di utenti con terze parti). Rispetto alla PSD2, le norme del Regolamento FIDA estendono i principi sull'accesso ai dati finanziari anche al di fuori del contesto dei conti di pagamento.

[292] Per esteso: Commissione europea, Proposta di Regolamento del Parlamento europeo e del Consiglio relativo a un quadro per l'accesso ai dati finanziari e che modifica i regolamenti (UE) n. 1093/2010, (UE) n. 1094/2010, (UE) n. 1095/2010 e (UE) 2022/2554, Bruxelles, 28.6.2023 COM(2023) 360 final 2023/0205 (COD), https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX%3A52023PC0360

Nell'aprile del 2024, il Parlamento europeo ha approvato la propria posizione sul testo della proposta della Commissione, mentre il Consiglio dell'UE ha approvato il suo orientamento a inizio dicembre scorso. Nel 2025, il Parlamento e il Consiglio hanno al momento avviato il processo di negoziazione con la partecipazione della Commissione (i cd. "triloghi") con l'obiettivo di giungere ad un testo concordato, per poi procedere all'approvazione formale del Regolamento.

V.4 RESPONSABILITÀ CIVILE - IL RITIRO DELLA PROPOSTA DI DIRETTIVA SULLA RESPONSABILITÀ DA IA

Nell'ultimo anno, la proposta di direttiva sulla responsabilità da intelligenza artificiale, presentata dalla Commissione il 28 settembre 2022 (COM(2022) 496), non ha conosciuto progressi significativi nell'*iter* di approvazione.

Così, a inizio 2025, la Commissione ha annunciato il ritiro della proposta nel suo programma di lavoro per il 2025, pubblicato il 12 febbraio 2025[293], in cui osserva come non vi sia: "nessun

[293] Cfr. Commissione europea, Comunicazione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Commission work programme 2025*, *Moving forward together: A Bolder*, *Simpler*, *Faster Union*, 11.2.2025 COM(2025) 45 final, si veda in particolare Allegato IV, "Withdrawals", punto 32, https://commission.europa.eu/strategy-and-policy/strategy-documents/commission-work-programme/commission-work-programme-2025_en?prefLang=it.

accordo prevedibile - e - la Commissione valuterà se presentare un'altra proposta o scegliere un altro tipo di approccio".

V.5 Focus sull'Italia

V.5.1 Il disegno di legge sull'IA – elementi fondamentali

In data 20 maggio 2024, il Governo italiano ha presentato un disegno di legge in materia di intelligenza artificiale[294], con il dichiarato scopo di operare un bilanciamento tra opportunità e rischi, prevedendo norme di principio e disposizioni di settore che, da un lato, promuovano l'utilizzo delle nuove tecnologie per il miglioramento delle condizioni di vita dei cittadini e, dall'altro, forniscano soluzioni per la gestione del rischio fondate su una visione antropocentrica.

Il panorama italiano si è poi arricchito anche del lavoro realizzato da un Comitato di 14 esperti di nomina governativa, volto a fornire aiuto nella definizione del quadro normativo nazionale in materia. Nel mese di luglio 2024, infatti, il gruppo di esperti ha elaborato e pubblicato la "Strategia Italiana per

[294] V. il testo del ddl originario, Atto Senato n. 1146, XIX Legislatura, Disposizioni e delega al Governo in materia di intelligenza artificiale: https://www.senato.it/leg/19/BGT/Schede/Ddliter/58262.htm.

l'Intelligenza Artificiale 2024-2026"[295]. Dopo aver analizzato il contesto internazionale e italiano, la Strategia individua quattro macroaree per gli obiettivi proposti (Strategia per la Ricerca, per la Pubblica Amministrazione, per le Imprese e per la Formazione) e definisce le modalità di attuazione, suggerendo anche l'istituzione di un nuovo soggetto (Fondazione per l'intelligenza artificiale) responsabile dell'attuazione, del coordinamento e del monitoraggio delle singole iniziative. Questa architettura istituzionale non è stata poi ripresa nel testo del disegno di legge, che però - come si vedrà - prevede l'aggiornamento annuale della strategia.

Il disegno di legge originario (atto n. 1146) è stato oggetto di un parere della Commissione europea nel mese di novembre 2024[296], in occasione del quale sono stati sollevati alcuni rilievi.

Accanto a osservazioni/indicazioni su: la necessità di un riferimento specifico al regolamento europeo sull'intelligenza

[295] DIPARTIMENTO PER LA TRASFORMAZIONE DIGITALE e AGID, *Strategia italiana per l'Intelligenza artificiale*, luglio 2024, documento scaricabile da qui: https://innovazione.gov.it/notizie/articoli/strategia-italiana-per-l-intelligenza-artificiale-2024-2026/.

[296] Commissione europea, parere circostanziato (C(2024) 7814) del 4 novembre 2024. I contenuti del parere sono disponibili nel resoconto istituzionale del Senato alla pagina: https://www.senato.it/japp/bgt/showdoc/19/SommComm/0/1435866/index.html?part=doc_dc-allegato_a:1.

artificiale nel testo e sulla necessità di coerenza definitoria rispetto all'AI Act, sulla la necessità di chiarire il concetto di dati "critici", sulla disciplina organica dei casi di uso di sistemi di intelligenza artificiale per finalità illecite, sulla previsione italiana originaria circa i fornitori di piattaforme per la condivisione di video soggetti alla giurisdizione italiana e relative misure a tutela del grande pubblico da contenuti informativi generati o modificati con l'IA.

Sono stati altresì evidenziati taluni altri rilievi in materia di *watermark*, in tema di impiego di sistemi di IA non ad alto rischio nelle professioni intellettuali, in materia di impieghi dei sistemi di IA nell'attività giudiziaria, in relazione agli obblighi informativi a carico degli operatori sanitari e in relazione alla designazione delle autorità nazionali competenti con lo stesso livello di indipendenza previsto 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.

Anche alla luce dei rilievi e osservazioni fatte dalla Commissione europea, il testo di disegno di legge ha subito talune revisioni rispetto alla versione originaria; l'articolato è stato approvato in Senato in prima lettura in data 20 marzo 2025 e, al momento in cui si scrive, il disegno di legge è ancora all'esame della Camera dei deputati[297]. Nel seguito si analizzano i principali contenuti del testo in esame alla Camera (Atto A.C. 2316)[298].

In coerenza agli obiettivi dichiarati dal Governo, il disegno di legge (nel seguito, per brevità "ddl") stabilisce al Capo I una serie di norme di principio (artt. 1-6) volte a garantire che tutte le attività all'interno del ciclo di vita dei sistemi di IA si svolgano con attenzione alla tutela dei diritti umani e alla salvaguardia dei principi che integrano lo Stato di diritto.

Sul piano tassonomico (art. 2) i sistemi e i modelli di IA sono definiti allo stesso modo di quanto fatto dalle istituzioni europee mediante l'AI Act attraverso un esplicito rinvio, rispettivamente, all'articolo 3, punto 1) e all'articolo 3, punto 63), del regolamento

[297] Il disegno di legge è all'esame delle Commissioni riunite nona e decima della Camera; la Commissione è previsto voti emendamenti e mandato al relatore per la seconda settimana di giugno 2025.

[298] Ad ogni modo, il testo trasmesso alla Camera (Atto Camera 2316) è disponibile qui: https://www.camera.it/leg19/126?&leg=19&idDocumento=2316.

(UE) 2024/1689. I principi generali che devono informare le attività dei sistemi e dei modelli di intelligenza artificiale per finalità generali sono indicati dall'art. 3 che evidenzia (art. 3, comma I) come la ricerca, la sperimentazione, lo sviluppo, l'adozione, l'applicazione e l'utilizzo di sistemi e di modelli di intelligenza artificiale per finalità generali avvengono nel rispetto dei diritti fondamentali e delle libertà previste dalla Costituzione, del diritto dell'Unione Europea e dei princìpi di trasparenza, proporzionalità, sicurezza, protezione dei dati personali, riservatezza, accuratezza, non discriminazione, parità dei sessi e sostenibilità.

Sempre quanto ai principi generali, il testo di ddl oggi all'esame della Camera chiarisce (art. 3, comma I) che lo sviluppo di sistemi e di modelli di intelligenza artificiale per finalità generali avviene su dati e tramite processi di cui devono essere garantite e vigilate la correttezza, l'attendibilità, la sicurezza, la qualità, l'appropriatezza e la trasparenza, secondo il principio di proporzionalità in relazione ai settori nei quali sono utilizzati. I sistemi e i modelli di intelligenza artificiale per finalità generali devono essere sviluppati e applicati nel rispetto dell'autonomia e del potere decisionale dell'uomo, della prevenzione del danno,

della conoscibilità, della trasparenza, della spiegabilità e dei princìpi di rispetto cui all'art. 3 comma 1, assicurando la sorveglianza e l'intervento umano.

Il testo, infine, precisa in modo esplicito (art. 3, comma 4) che l'utilizzo di sistemi di intelligenza artificiale non deve pregiudicare lo svolgimento con metodo democratico della vita istituzionale e politica e l'esercizio delle competenze e funzioni delle istituzioni territoriali sulla base dei princìpi di autonomia e sussidiarietà e che (art. 3, comma 6), per garantire il rispetto dei diritti e dei princìpi di cui all'art. 3, deve essere assicurata - quale precondizione essenziale - la cybersicurezza lungo tutto il ciclo di vita dei sistemi e dei modelli di intelligenza artificiale per finalità generali, secondo un approccio proporzionale e basato sul rischio, nonché l'adozione di specifici controlli di sicurezza, anche al fine di assicurarne la resilienza contro tentativi di alterarne l'utilizzo, il comportamento previsto, le prestazioni o le impostazioni di sicurezza.

Sempre al Capo I, il ddl prosegue con le disposizioni relative ai princìpi in materia di informazione e riservatezza dei dati personali

(art. 4), principi in materia di sviluppo economico (art. 5) e con le disposizioni in materia di sicurezza e difesa nazionale (art. 6).

Il Capo II (artt. 7-18) del testo oggi alla Camera concerne disposizioni specifiche per alcuni settori come in ambito sanitario e disabilità (art. 7 - 10), il lavoro (art. 11 e 12) le professioni intellettuali (art. 13), la pubblica amministrazione (art. 14) e l'attività giudiziaria (art. 15) al fine di garantire che l'utilizzo dell'IA sia praticato non solo per sfruttare tutte le opportunità legate alle nuove tecnologie, ma anche sulla base di criteri che assicurino la gestione e la soluzione dei rischi connessi allo svolgimento di tali attività.

A titolo esemplificativo l'art. 7 - concernente l'uso dell'IA in ambito sanitario - precisa al comma I che tale utilizzo non possa selezionare o condizionare l'accesso alle prestazioni sanitarie in maniera discriminatoria, garantendo in ogni caso il diritto dell'interessato a essere informato circa l'utilizzo di tali tecnologie. Al comma 5 dell'art. 7 viene previsto che i sistemi di IA in ambito sanitario costituiscono un aiuto nei processi di prevenzione, diagnosi, cura e scelta terapeutica che, pertanto,

lasciano impregiudicata la decisione rimessa sempre al personale sanitario.

L'art. 11 del testo oggi in esame alla Camera applica il principio antropocentrico all'utilizzo dell'IA nel mondo del lavoro, chiarendo che l'intelligenza artificiale è impiegata per migliorare le condizioni di lavoro, tutelare l'integrità psicofisica dei lavoratori, accrescere la qualità delle prestazioni lavorative e la produttività delle persone in conformità al diritto dell'Unione Europea; allo stesso tempo (art. 11, comma II) ogni utilizzo dell'IA in ambito lavorativo deve essere sicuro, affidabile e trasparente, essendo tenuto il datore di lavoro ad informare il lavoratore dell'utilizzo di sistemi di intelligenza artificiale. Al comma III dell'art. 11, infine, si prevede che l'intelligenza artificiale nell'organizzazione e nella gestione del rapporto di lavoro garantisca l'osservanza dei diritti inviolabili del lavoratore, senza discriminazioni in funzione del sesso, dell'età, delle origini etniche, del credo religioso, dell'orientamento sessuale, delle opinioni politiche e delle condizioni personali, sociali ed economiche, in conformità con il diritto dell'Unione.

Sempre in connessione con il mondo del lavoro, l'art. 12 prevede la costituzione, presso il Ministero del Lavoro e delle Politiche Sociali, dell'Osservatorio sull'adozione di sistemi di intelligenza artificiale nel mondo del lavoro, con il compito di definire una strategia sull'utilizzo dell'intelligenza artificiale in ambito lavorativo, monitorare l'impatto sul mercato del lavoro e identificare i settori lavorativi maggiormente interessati dall'avvento dell'intelligenza artificiale.

All'art. 12 - con riferimento alle professioni intellettuali - si statuisce che l'utilizzo di sistemi di intelligenza artificiale è finalizzato al solo esercizio delle attività strumentali e di sostegno all'attività professionale e con prevalenza del lavoro intellettuale oggetto della prestazione d'opera (comma I). Per assicurare il rapporto fiduciario tra professionista e cliente, le informazioni relative ai sistemi utilizzati dal professionista sono comunicate al soggetto destinatario della prestazione intellettuale con linguaggio chiaro, semplice ed esaustivo (comma II).

Dopo l'articolo 14 dedicato alla PA, il ddl passa a disciplinare (art. 15) l'uso dell'IA nell'ambito dell'attività giudiziaria, stabilendo (comma I) che nell'utilizzo dell'IA nell'amministrazione della giustizia è sempre riservata al magistrato ogni decisione sull'interpretazione e sull'applicazione della legge, sulla valutazione dei fatti e delle prove e sull'adozione dei provvedimenti. Sarà il Ministero della Giustizia a disciplinare gli impieghi dei sistemi di intelligenza artificiale per l'organizzazione dei servizi relativi alla giustizia, per la semplificazione del lavoro giudiziario e per le attività amministrative accessorie (comma II). Secondo il comma II del medesimo articolo, fino alla compiuta attuazione del regolamento (UE) 2024/1689, la sperimentazione e l'impiego dei sistemi di intelligenza artificiale negli uffici giudiziari ordinari sono autorizzati dal Ministero della Giustizia, sentite le autorità nazionali indicate nel ddl all'art. 20.

Il Capo III (artt. 19-24) del ddl è dedicato alla definizione della strategia nazionale e della *governance* dell'IA. L'art. 19 dispone che la strategia è predisposta e aggiornata dalla competente struttura presso la Presidenza del Consiglio dei Ministri, d'intesa

con le Autorità nazionali per l'intelligenza artificiale di cui all'art. 20, sentiti il Ministro delle Imprese e del *Made in Italy*, il Ministro dell'Università e della ricerca e il Ministro della Difesa sentito il parere dei ministeri competenti *ratione materiae* (ad es., quello delle imprese e del *Made in Italy* per i profili di politica industriale e di incentivazione) ed è approvata con cadenza almeno biennale dal Comitato interministeriale per la transizione digitale (CITD).

Ai sensi dell'art. 20 sono designate come autorità nazionali per l'IA: l'Agenzia per l'Italia digitale (AgID) e l'Agenzia per la cybersicurezza nazionale (ACN), incaricate di garantire l'applicazione e l'attuazione della normativa nazionale e dell'Unione europea in materia di IA (ferma restando l'attribuzione a Banca d'Italia, CONSOB e IVASS del ruolo di autorità di vigilanza del mercato).

L'AgID sarà responsabile della promozione dell'innovazione e dello sviluppo dell'intelligenza artificiale, della definizione delle procedure e dell'esercizio delle funzioni e dei compiti in materia

di notifica, valutazione, accreditamento e monitoraggio dei soggetti incaricati di verificare la conformità dei sistemi di IA.

L'ACN sarà responsabile della vigilanza, incluse le attività ispettive e sanzionatorie, dei sistemi di IA, nonché per la promozione e lo sviluppo dell'intelligenza artificiale relativamente ai profili di cybersicurezza.

Entrambe le Autorità, per quanto di rispettiva spettanza, assicurano l'istituzione e la gestione congiunta di spazi di sperimentazione finalizzati alla realizzazione di sistemi di intelligenza artificiale conformi alla normativa nazionale e dell'Unione Europea, sentiti il Ministero della Difesa e il Ministero della Giustizia per gli ambiti di competenza (resta salvo e fermo quanto previsto dall'articolo 36, commi da 2-bis a 2-*novies*, del decreto legge 30 aprile 2019, n. 34, convertito, con modifiche, dalla Legge 28 giugno 2019, n. 58, per quanto concerne la sperimentazione di sistemi di intelligenza artificiale destinati ad essere immessi sul mercato, messi in servizio o usati da istituti finanziari).

Infine, ferma restando l'attribuzione a Banca d'Italia, CONSOB e IVASS del ruolo di vigilanza dei mercati ai sensi della disciplina vigente, l'AgID è designata quale autorità di notifica ai sensi dell'articolo 70 dell'AI Act, mentre l'ACN è designata quale autorità di vigilanza del mercato e punto di contatto unico con le istituzioni dell'Unione Europea ai sensi del medesimo articolo 70 (art. 20, comma II).

Il Capo IV (art. 25) contiene norme concernenti la tutela degli utenti, preoccupandosi di introdurre norme specifiche in materia di diritto d'autore, attraverso la modifica dell'art. 1 della Legge sul Diritto d'Autore (L. 22 aprile 1941, n. 633) e ampliando le tutele ivi previste anche alle opere dell'ingegno umano di carattere creativo: «anche laddove create con l'ausilio di strumenti di intelligenza artificiale, purché costituenti il risultato del lavoro intellettuale dell'autore». Inoltre, il ddl stabilisce un'ulteriore modifica dell'anzidetta Legge con la previsione dell'inserimento in essa di un nuovo ulteriore articolo, il 70-*septies*, relativo a riproduzioni ed estrazioni da opere o da altri materiali contenuti in rete o in banche di dati a cui si ha legittimamente accesso, ai fini

dell'estrazione di testo e di dati attraverso modelli e sistemi di intelligenza artificiale, anche generativa.

Il Capo V (art. 26), si concentra sui riflessi penalistici dell'utilizzo degli strumenti di IA, introducendo un aggravante comune all'art. 61, co. 1, n. 11 *decies* c.p. per i fatti commessi mediante l'impiego di sistemi di IA quando gli stessi, per la natura o per le modalità di utilizzo, abbiano costituito mezzo insidioso ovvero quando il loro impiego abbia ostacolato la pubblica o privata difesa o aggravato le conseguenze del reato.

Inoltre, sempre alla lett. b) del I comma dell'art. 26 del ddl è prevista l'aggiunta di un comma all'art. 294 del c.p. (Attentati contro i diritti politici del cittadino), prevedendo la pena della reclusione da due a sei anni se l'inganno è posto in essere mediante l'impiego di sistemi di intelligenza artificiale. Nel ddl, inoltre, è prevista l'introduzione di una nuova fattispecie di reato: dopo l'art. 612-*quater* c.p., è proposto l'inserimento di un nuovo articolo (art. 612-*quater*) che prevedere il delitto di illecita diffusione di contenuti generati o manipolati artificialmente.

Quanto al Codice Civile, il ddl prevede l'inserimento di un nuovo comma all'art. 2637 c.c. in relazione all'aggiotaggio, che stabilisce la reclusione da due a sette anni se il fatto è commesso mediante l'impiego di sistemi di intelligenza artificiale. Quanto poi agli interventi sulla normativa speciale, nel testo oggi all'esame della Camera ne sono indicati alcuni modificativi sulla Legge sul Diritto d'Autore e sul Testo Unico delle disposizioni in materia di intermediazione finanziaria (in particolare, sull'art. art.185 TUF).

Il ddl si chiude con il Capo VI (art. 27-28) relativo alle disposizioni finanziarie e finali.

V.5.2 IA GENERATIVA E PROTEZIONE DATI: LE DECISIONI DEL GARANTE ITALIANO SU CHATGPT E DEEPSEEK

Il Garante per la Protezione dei Dati Personali italiano si è distinto a livello europeo per le sue decisioni nei confronti dei sistemi *chatbot* basati sui modelli di intelligenza artificiale generativa più avanzati e diffusi.

Gli ultimi aggiornamenti in questo senso consistono nel Provvedimento 2 novembre 2024[299] di chiusura dell'istruttoria nei confronti della società americana OpenAI per il "servizio ChatGPT" e il Provvedimento 30 gennaio 2025[300] di "blocco" (tecnicamente, limitazione del trattamento di dati personali) della cinese DeepSeek sul territorio italiano.

Con il primo provvedimento, il GPDP italiano ha chiuso l'istruttoria avviata nel marzo 2023 nei confronti della società OpenAI in relazione al trattamento illecito di dati personali tramite il servizio di intelligenza artificiale generativa ChatGPT. La decisione del Garante sanziona OpenAI per violazioni nella gestione del servizio, avendo accertato, in particolare, come la società non abbia notificato una violazione dei dati avvenuta nel periodo di indagine, abbia trattato dati personali senza una base giuridica adeguata, abbia violato il principio di trasparenza e gli obblighi informativi nei confronti degli utenti previsti dal GDPR

[299] GARANTE PER LA PROTEZIONE DEI DATI PERSONALI, Provvedimento del 2 novembre 2024, Registro dei provvedimenti n. 755 del 2 novembre 2024 [doc. web n. 10085455], disponibile alla pagina https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10085455.

[300] GARANTE PER LA PROTEZIONE DEI DATI PERSONALI, Provvedimento del 30 gennaio 2025, Registro dei provvedimenti n. 33 del 30 gennaio 2025 [doc. web n. 10098477], disponibile alla pagina: https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/10098477.

e non abbia adottato sistemi di verifica dell'età, esponendo i minori a contenuti inappropriati.

Come misura per garantire maggiore trasparenza, il provvedimento prevede una misura basata sul nuovo potere a disposizione del Garante introdotto all'articolo 166, comma 7 del Codice *Privacy* italiano (d.lgs. n. 196/2003): OpenAI ha previsto di condurre una campagna informativa di sei mesi su vari media, spiegando il funzionamento di ChatGPT e i diritti degli utenti riguardo ai propri dati, tra cui opposizione, rettifica e cancellazione. L'obiettivo è sensibilizzare il pubblico sulla possibilità di escludere i propri dati dall'addestramento dell'IA.

Il Garante ha inoltre imposto a OpenAI una sanzione di quindici milioni di euro, che tiene conto comunque dell'atteggiamento collaborativo della società durante l'istruttoria. Nel frattempo, OpenAI ha trasferito il proprio quartier generale europeo in Irlanda. Di conseguenza, trova applicazione il principio del cd. "*one stop shop*" previsto dagli articoli 56 e 60 del GDPR, secondo cui l'autorità dello Stato in cui l'impresa titolare (o responsabile)

del trattamento è stabilita, è competente, quale autorità di controllo "capofila" anche per i trattamenti transfrontalieri di tale impresa. Il Garante italiano ha perciò trasmesso gli atti all'Autorità irlandese per la protezione dei dati (DPC). Quest'ultima diventerà l'ente di riferimento per eventuali ulteriori indagini sulle violazioni che potrebbero persistere dopo l'apertura della sede europea della società OpenAI.

Con il Provvedimento del 30 gennaio 2025, invece, il Garante ha ordinato in via d'urgenza - ai sensi dell'art. 58, par. 2, lett. f) del GDPR – la: "limitazione definitiva del trattamento di dati personali" di interessati che si trovano nel territorio italiano in relazione al "servizio DeepSeek". Le destinatarie della misura sono le società cinesi che gestiscono il servizio (Hangzhou DeepSeek Artificial Intelligence Co., Ltd., e Beijing DeepSeek Artificial Intelligence Co., Ltd) che avevano risposto alla precedente richiesta di informazioni del Garante italiano, affermando di non operare in Italia.

L'ordine - che dovrebbe comportare sostanzialmente il blocco del servizio DeepSeek sul territorio italiano - si fonda anche in questo caso sulla mancata dimostrazione di una valida base giuridica per il trattamento dei dati personali, sulla carenza dell'informativa presentata agli utenti del servizio, oltre che sulla circostanza della conservazione dei dati personali nel territorio della Repubblica popolare cinese senza le misure di sicurezza del trattamento imposte dal GDPR.

Le misure del Garante italiano mostrano chiaramente i conflitti (potenziali e attuali) tra l'utilizzo di tecnologie di intelligenza artificiale avanzate e la tutela di diritti fondamentali come quello alla *privacy* e alla protezione dei dati personali. Le modalità con cui sono utilizzati ed eventualmente "diffusi" dati personali da parte dei sistemi di intelligenza artificiale generativa è oggetto di attenzioni e richiede probabilmente ancora degli adeguamenti, sia dal punto di vista tecnologico che giuridico.

A tal proposito, è da rilevare come alle iniziative del Garante italiano non si sono aggiunte quelle di altre autorità per la

protezione dei dati a livello europeo: sarebbe auspicabile un'azione coordinata da parte delle diverse autorità nazionali, sulla base delle modalità di cooperazione previste dal GDPR e delle indicazioni eventualmente fornite dal Comitato e dal Garante europeo per la protezione dei dati.

BOX 7. EVOLUZIONE DEL DIRITTO D'AUTORE NELL'ERA DELL'IA.

L'evoluzione delle macchine computazionali ha avuto uno sviluppo esponenziale negli ultimi anni, rendendo i sistemi di IA in grado non solo di assistere l'uomo nelle sue mansioni, ma anche di compiere in modo "indipendente" compiti creativi simili a quelli umani.

Se da un lato l'IA si è rivelata una scoperta provvidenziale per l'esecuzione di compiti caratterizzati da una profonda complessità; dall'altro lato, essa ha sollevato molti dubbi rispetto al funzionamento della disciplina autoriale.

A tal proposito, sebbene l'AI Act non affronti estesamente il problema della tutela del diritto d'autore, obbliga i fornitori di modelli di IA per finalità generali ad attuare una politica che risponda al diritto dell'Unione. in tale materia. Gli obblighi in questione sono anche declinati all'interno del codice di condotta (cfr. *supra* par.V.2.2), che contiene una sezione specifica dedicata al rispetto del diritto d'autore rispetto alle attività di *web crawling,* finalizzata all'addestramento dei modelli e al rischio di generazione di *output* lesivi di diritti altrui.

Sul piano nazionale, il diritto d'autore italiano ruota da sempre intorno a una concezione antropocentrica, che riconosce come opere dell'ingegno esclusivamente quei prodotti frutto della creatività umana. Cosicché, nessuna tutela potrebbe essere offerta alle opere generate "autonomamente" dai sistemi di intelligenza artificiale.

Questo ragionamento si fonda sul concetto di "creatività qualificata", secondo cui un'opera dell'ingegno, per poter essere considerata tale, deve non solo essere frutto dell'attività di un autore (umano), ma anche: «riflettere la sua personalità e le sue scelte creative». Senonché, a dover essere condivisa è quella impostazione che sostiene la teoria della "creatività semplice", che si fonda sul presupposto che un'opera, per poter essere qualificata come creazione dell'ingegno deve: «contenere forme espressive scelte in maniera discrezionale tra più varianti equipollenti», indipendentemente dal fatto che tali scelte siano state effettuate da un essere umano oppure da una macchina computazionale.

Del resto, chiara espressione dell'accoglimento di tale teoria era il nuovo art. 1 della legge italiana sul diritto d'autore (l. 22 aprile 1941, n. 633[301]) così come modificato dall'art. 25 del d.d.l. italiano sull'IA, il quale ha ampliato le tutele ivi previste alle opere dell'ingegno umano di carattere creativo: «anche là dove create con l'ausilio di strumenti di intelligenza artificiale, purché costituenti il risultato del lavoro intellettuale dell'autore».

Una volta accertata l'applicabilità della disciplina autoriale anche alle creazioni prodotte dai sistemi di intelligenza artificiale, è indispensabile interrogarsi su chi sia il soggetto titolare dei diritti morali e patrimoniali d'autore che insistono su queste opere. Le opzioni che sono state proposte in

[301] Legge 22 aprile 1941, n. 633, Protezione del diritto d'autore e di altri diritti connessi al suo esercizio, cfr. https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1941-04-22;633!vig=2025-05-27.

dottrina sono tre: 1.) attribuire la paternità dell'opera alla intelligenza artificiale[302]; 2.) considerare titolare del diritto d'autore il programmatore o sviluppatore dell'IA[303]; 3.) riconoscere in capo all'utilizzatore finale la titolarità del diritto d'autore[304]. La posizione che sembra essere quella maggiormente sostenibile è l'ultima, che si fonda su una visione innovativa della figura d'autore, considerato non più solo come persona fisica che dà origine al prodotto dell'ingegno, ma anche come individuo che assume l'iniziativa funzionale alla realizzazione dell'opera.

Infine, un'altra sfida che l'utilizzo dei sistemi di IA pone è la regolamentazione della fase di addestramento delle macchine computazionali che utilizzano opere protette dal diritto d'autore. Tale questione si è posta nel caso New York Times Company v. OpenAI (United States District Court, Southern District of New York, Civil Action No. 1:23 cv 11195) dove, il 27 dicembre 2023, il *New York Times* ha citato in giudizio Microsoft Corporation e OpenAI invocando una lesione dei diritti di *copyright* a causa dell'utilizzo illecito di articoli e contenuti della testata giornalistica per il *training* di modelli di IA generativa.

[302] P. GITTO, "*New York Times* v. OpenAI, Microsoft *et al*.: conflitti attuali fra intelligenza artificiale e diritto d'autore", *Giustizia Civile*, 2/2024, p. 185 ss., https://giustiziacivile.com/societa-e-concorrenza/approfondimenti/new-york-times-vs-openai-microsoft-et-al-conflitti-attuali-fra); G. FROSIO, "L'(I)Autore inesistente: una tesi tecno-giuridica contro la tutela dell'opera generata dall'Intelligenza Artificiale", *Annali Italiani di Diritto d'Autore* (*AIDA*), Vol. 29, 2020, pp. 52-91.

[303] F. FERRETTI, "Intelligenza artificiale e diritto d'autore: quale tutela per il robot creatore?", *Rassegna del diritto della moda e delle arti*, 2022, pp. 68-106; E. BONADIO, L. MCDONAGH, "Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity", *Intellectual Property Quarterly*, 2020, pp. 112-137; S. GUIZZARDI, "La protezione d'autore dell'opera dell'ingegno creata dall'Intelligenza Artificiale", *Annali Italiani di Diritto d'Autore* (*AIDA*), XXVII, 2028, pp.1-31.

[304] M. D'ONOFRIO, "Azioni e creazioni dell'intelligenza artificiale", *Tecnologie e diritto*, 1/2024; C. DI BENEDETTO, "Creatività artificiale e diritti d'autore", *Annali Italiani di Diritto d'Autore (AIDA*), Vol.31, 2022, pp.388-402.

L'accusa riguardava l'utilizzo di tali contenuti senza alcun accordo tra le parti e, di conseguenza, senza il consenso dei titolari del diritto d'autore. In particolare, Microsoft e OpenAI sono stati accusati di "*willful infringement*", ovvero violazione volontaria del diritto d'autore, motivo per cui la New York Times Company riteneva che essi fossero a conoscenza o avrebbero dovuto conoscere le possibili conseguenze dell'utilizzo (illecito) di contenuti coperti da *copyright*[305].

Dall'altro lato OpenAI, nel tentativo di giustificare il proprio contegno, invoca il c.d. *fair use*, ovvero una eccezione al diritto d'autore contenuta nell'art. 107 del *Copyright Act* del 1976 che consente, in determinate circostanze, l'utilizzo limitato di materiale protetto da *copyright* senza il previo consenso del titolare del diritto in questione. La controversia risultava - all'inizio del 2025 - nella fase della *discovery*, ossia della produzione in giudizio di documenti per l'istruttoria, il cui termine era previsto per fine aprile e in cui ad OpenAI è stato ordinato di mostrare alle controparti i dati di addestramento (*training data*) utilizzati per sviluppare i propri modelli.

[305] Gli atti fondamentali della controversia possono essere consultati alla pagina: https://www.bakerlaw.com/new-york-times-v-microsoft/. Da aprile 2025 la causa del *New York Times* è trattata unitamente a diverse altre cause di editori con oggetto comune. Gli atti fondamentali sono consultabili alla pagina: www.bakerlaw.com/in-re-openai-inc-copyright-infringement-litigation/. Per gli ultimi sviluppi v. anche https://tmsnrt.rs/4ljj2PO.

BOX 8. L'UTILIZZO DELL'INTELLIGENZA ARTIFICIALE DA PARTE DELLA PUBBLICA AMMINISTRAZIONE

L'utilizzo dell'intelligenza artificiale nella Pubblica Amministrazione italiana è già una realtà, con *best practice* e casi d'uso che sono descritti nel Piano Triennale per l'Informatica nella PA di AgID[306]. D'altra parte, l'utilizzo dell'intelligenza artificiale nella pubblica amministrazione è visto come un possibile strumento per aumentare l'efficacia e l'efficienza della stessa, tanto da risultare uno specifico capitolo della strategia sull'intelligenza artificiale italiana (v. *supra* par. V.5.1). In tale contesto, a inizio 2025 AgID ha posto in consultazione le linee guida per l'impiego di sistemi di intelligenza artificiale nel contesto pubblico[307].

Ciò nonostante, l'utilizzo delle tecnologie di IA nel contesto pubblico appare ancora frammentato, disomogeneo e spesso poco trasparente. In molti casi, le soluzioni adottate - sia sviluppate internamente che acquisite da fornitori esterni – sono messe in opera in assenza di riferimenti normativi espliciti o senza una reale supervisione pubblica dei modelli algoritmici sottostanti. Tale situazione è ulteriormente aggravata dalla mancanza di una cultura condivisa dell'uso dell'IA nella PA e dall'assenza di un quadro istituzionale pienamente conforme ai principi europei di indipendenza e *accountability*.

In particolare, si evidenzia la necessità di una ricognizione sistematica e aggiornata sull'adozione effettiva dell'IA da parte delle amministrazioni centrali e locali, anche in relazione agli obblighi previsti dal Codice dei

[306] Cfr. ultimo aggiornamento *Piano Triennale per l'informatica nella PA 2024-2026*, Edizione 2024-2026 (Aggiornamento 2025), https://www.agid.gov.it/it/agenzia/piano-triennale.

[307] Cfr. determinazione n. 17 del 17 febbraio 2025 - Avvio dell'*iter* di consultazione e informazione delle Linee guida per l'adozione dell'Intelligenza Artificiale nella Pubblica Amministrazione, https://www.agid.gov.it/it/notizie/intelligenza-artificiale-in-consultazione-le-linee-guida-pa.

Contratti Pubblici e alle linee guida AgID. L'attuale assetto di *governance*, con AgID e ACN operanti sotto la Presidenza del Consiglio dei ministri, potrebbe infatti non garantire il grado di indipendenza richiesto dal Regolamento europeo sull'IA.

In questo contesto, risulta urgente potenziare la trasparenza, la tracciabilità delle decisioni automatizzate e la formazione tecnica delle amministrazioni. Solo attraverso un rafforzamento delle competenze interne e una supervisione indipendente sarà possibile garantire un'adozione dell'IA nella PA che sia efficace, eticamente fondata e conforme alle norme europee.

Alcuni scenari estremi già ventilati in ambito pubblico, come la proposta di sostituire parte dei funzionari non eletti (i cosiddetti *civil servant*) con algoritmi decisionali, sollevano questioni ancora più profonde. Se da un lato, ciò potrebbe migliorare l'efficienza e la trasparenza amministrativa, dall'altro, rischia di indebolire i fondamenti democratici delle istituzioni, disancorando il processo decisionale dalla responsabilità politica e dal controllo umano. Il dibattito sulla *governance* algoritmica deve perciò affrontare anche il nodo della delega non elettiva del potere.

Di recente, il Cancelliere britannico Rachel Reeves ha annunciato un piano per sostituire almeno diecimila posti di lavoro nel servizio civile con l'intelligenza artificiale, come parte di un'iniziativa di riduzione dei costi annuale di due miliardi di sterline. Il piano prevede tagli del 15% ai budget amministrativi entro la fine del decennio, con l'IA che assumerà molti ruoli di "*back-office*", inclusi risorse umane, comunicazioni, politiche, approvvigionamento e gestione dell'ufficio.

Inoltre, un rapporto del Tony Blair Institute (https://institute.global/) suggerisce che oltre il 40% delle attività svolte dai lavoratori del settore

pubblico potrebbe essere parzialmente automatizzato tramite software basati sull'IA, come modelli di apprendimento automatico e modelli linguistici di grandi dimensioni, nonché *hardware* abilitato all'IA.

Sono iniziative che sollevano importanti questioni etiche e di *governance*, in particolare riguardo alla responsabilità, alla trasparenza e all'impatto sull'occupazione nel settore pubblico.

VII.

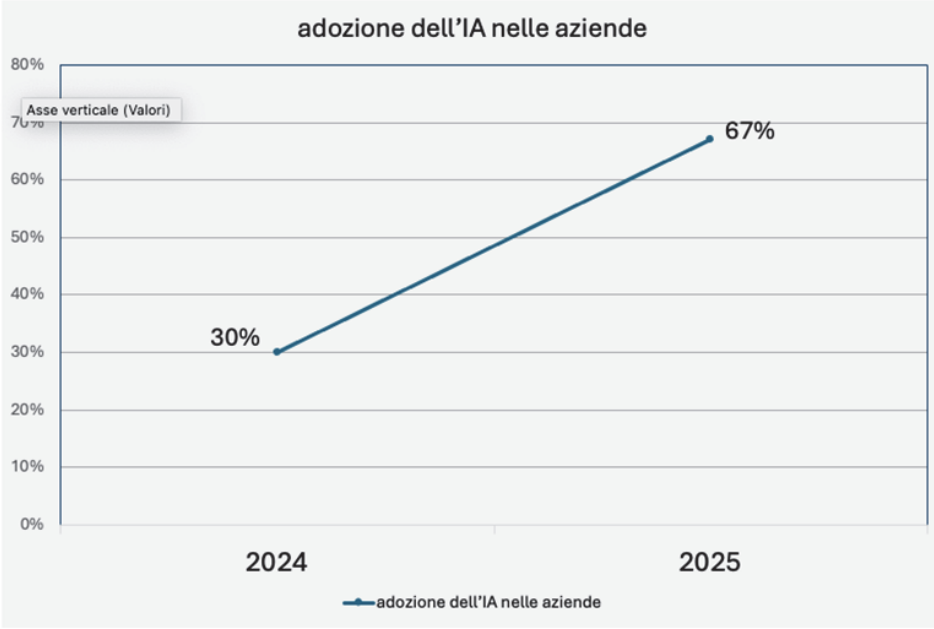
RISULTATI DEL QUESTIONARIO

Ad accompagnare il rapporto, così come nella prima edizione, è stata realizzata da Aspen Institute Italia una consultazione empirica, non rappresentativa in senso statistico, rivolta a un *target* qualificato di aziende italiane. Ciò ha permesso di tracciare una fotografia aggiornata, seppur non campionaria, di percezioni, strategie e dinamiche in atto nel mondo produttivo.

I dati che seguono, relativi all'edizione 2025, confermano il forte consolidamento dell'intelligenza artificiale nel tessuto industriale italiano. Rispetto all'anno precedente, emerge un significativo aumento delle aziende che dichiarano di aver già avviato iniziative concrete in questo ambito: si passa infatti dal 30% del 2024 al 67% del 2025.

Si tratta di un'evoluzione che non riguarda solo l'adozione tecnologica, ma anche la visione strategica e la consapevolezza dell'impatto sistemico dell'IA. Le imprese stanno ampliando il perimetro delle funzioni coinvolte, passando dalla sperimentazione nel *marketing* e nella *customer experience* a un'integrazione più strutturale nei processi decisionali, produttivi e organizzativi.

Il grafico che segue sintetizza il trend di crescita dell'adozione dell'IA tra il 2024 e il 2025. A seguire si presentano nel dettaglio le evidenze emerse dalla consultazione, arricchite da commenti e spunti interpretativi utili per orientare le scelte future.



RISULTATI*

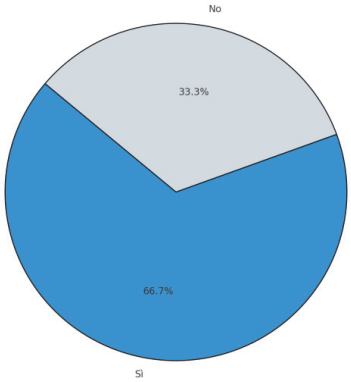
Come detto, nel corso di stesura del Rapporto è stato sottoposto un questionario a un *panel* di intervistati con l'obiettivo di comprendere il grado di adozione dell'Intelligenza Artificiale, i benefici percepiti, le sfide operative, nonché le aspettative future in termini di investimenti, formazione e sostegno normativo all'interno del settore industriale.

Il campione intervistato è stato di 54 Aziende - inizialmente suddiviso in tre macrocategorie - in seguito accorpato con l'obiettivo di fornire al lettore una vista d'insieme delle risposte pervenute sui singoli quesiti.

Si riporta qui di seguito una sintesi delle principali evidenze.

**Cura del questionario, Monica Coppi, Aspen Institute Italia.*

. La sua azienda nel corso dell'ultimo anno ha adottato soluzioni basate sull'Intelligenza Artificiale nei suoi processi operativi o decisionali?

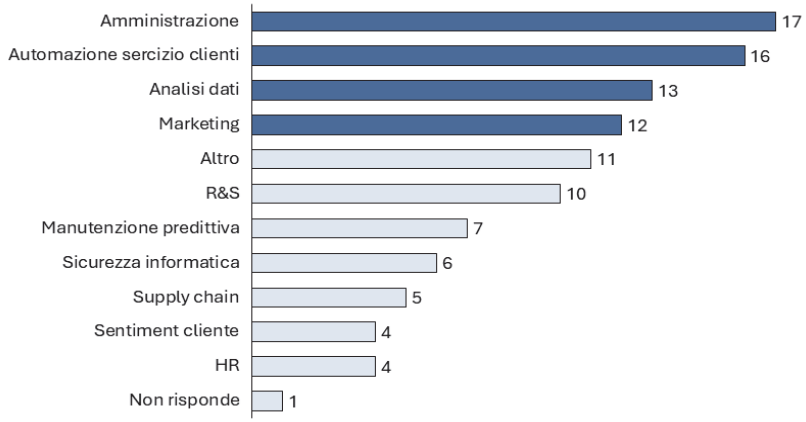


36 aziende delle 54 intervistate (pari al **66,7%**) hanno adottato nell'ultimo anno soluzioni basate sull'IA nei loro processi operativi o decisionali.

Questo dato suggerisce una tendenza di adozione significativa tra le aziende del campione, riflettendo una crescente maturità rispetto alla sperimentazione e all'integrazione dell'IA nei processi aziendali.

Si rileva inoltre un perfetto equilibrio tra **l'adozione di soluzioni esistenti (18 risposte) e lo sviluppo interno (18 risposte)**, segno di una duplice strategia: da un lato, rapidità nell'adozione attraverso tecnologie disponibili; dall'altro, capacità (o necessità) di personalizzazione attraverso sviluppo proprietario.

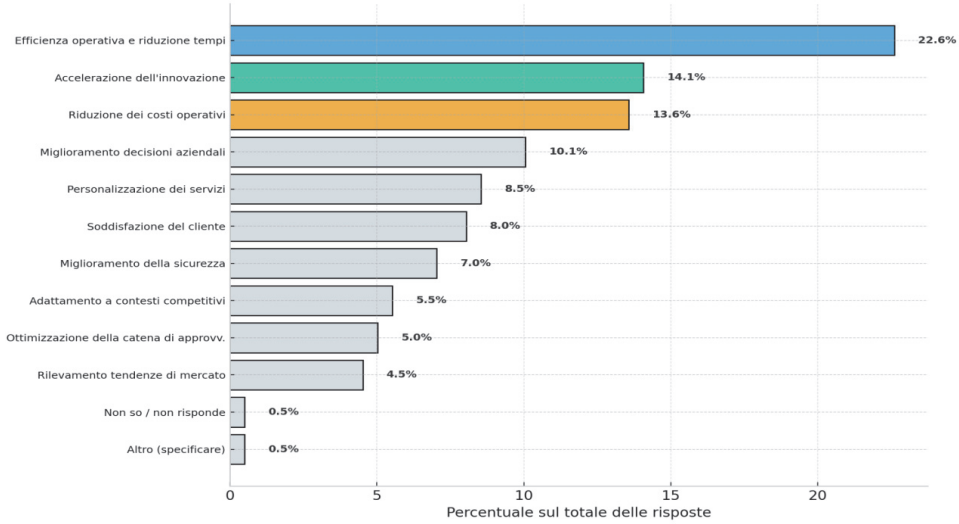
. In quali processi aziendali è stata utilizzata/implementata? *(possibili più risposte)*



L'**automazione dei processi amministrativi** (17 risposte) e quella del **servizio clienti** (16) sono gli ambiti principali di applicazione, seguite da **analisi dati** (13) e *marketing* (12).

Anche aree complesse come **ricerca e sviluppo** (10) e **manutenzione predittiva** (7) risultano già interessate dall'uso dell'IA.

. Quali benefici strategici ritiene che siano derivati o possano derivare per la sua azienda dall'utilizzo/implementazione dell'Intelligenza Artificiale? *(possibili più risposte)*

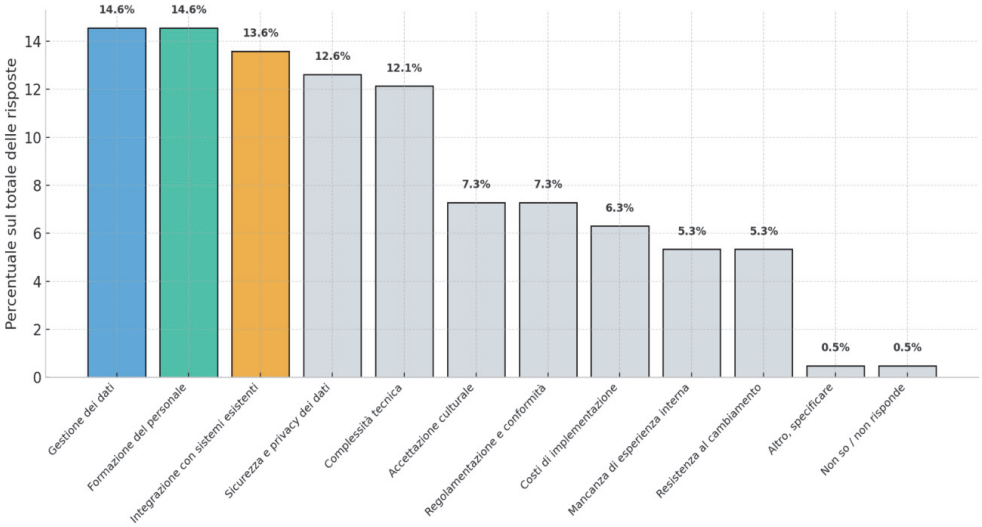


Le risposte al quesito mostrano una netta predominanza di alcuni benefici percepiti.

"**Efficienza operativa e riduzione dei tempi** (22,6% delle risposte)" è il beneficio più largamente riconosciuto. Ciò conferma l'associazione immediata tra IA e automazione, snellimento dei processi e riduzione dei tempi di esecuzione.

Anche "**Accelerazione dell'innovazione** (14,1%)" e "**Riduzione dei costi** (13,6%)" emergono come leve chiave per aumentare la competitività.

. Quali sfide prevede che la sua azienda potrebbe affrontare nell'implementare soluzioni basate sull'Intelligenza Artificiale?



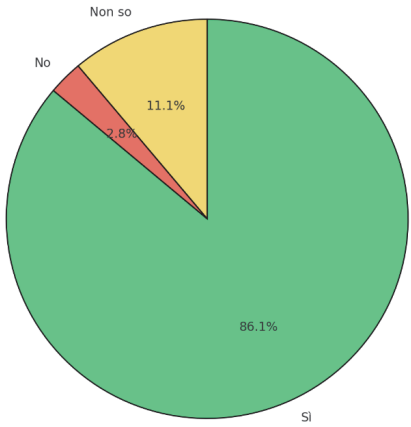
Le principali sfide individuate dalle aziende riguardano:

- **Formazione del personale e gestione dei dati (entrambe con il 14,6% delle risposte),** che si confermano come le barriere più critiche.

- **Segue da vicino l'integrazione con i sistemi esistenti (13,6%),** indicatore della difficoltà di adattare l'IA a infrastrutture *legacy*.

Questi dati indicano chiaramente che l'efficace messa in opera dell'IA non è solo una questione di tecnologia, ma richiede una solida infrastruttura informativa, competenze adeguate e compatibilità con i sistemi esistenti.

. La sua azienda ha intenzione di estendere l'utilizzo dell'Intelligenza Artificiale in altre aree aziendali?

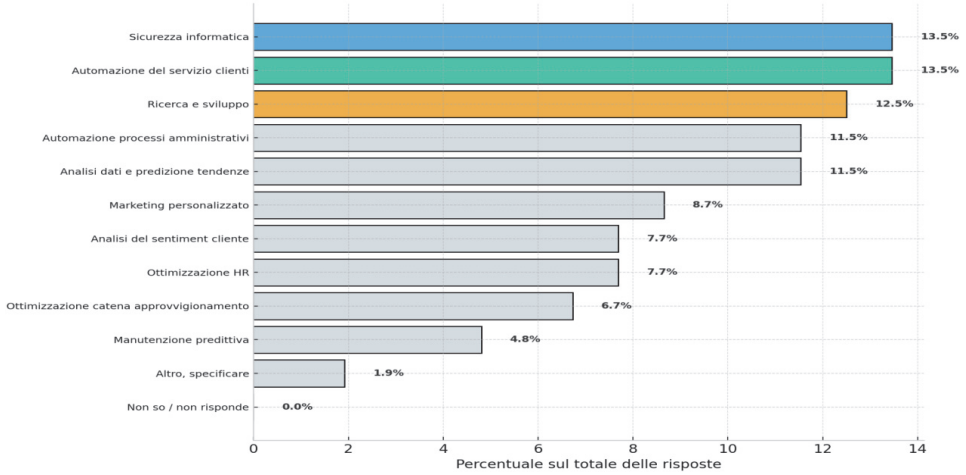


L'analisi delle risposte mostra una **forte tendenza all'estensione dell'Intelligenza Artificiale** in nuove aree aziendali:

- Ben **l'86,1%** delle aziende ha risposto **positivamente** all'intenzione di estendere l'uso dell'IA.
- Solo il **2,8%** ha dichiarato esplicitamente di **non avere questa intenzione**.
- L'**11,1%** si colloca in una **posizione interlocutoria**, indicando che non è ancora stata presa una decisione definitiva.

Questo dato evidenzia una **tendenza marcata verso una diffusione più ampia e trasversale dell'IA nei processi aziendali**, sintomo di fiducia negli effetti positivi finora sperimentati e nella maturazione del contesto tecnologico.

. Quali aree sono attualmente considerate per questa espansione?



Le aziende che intendono estendere l'utilizzo dell'Intelligenza Artificiale si concentrano su alcune aree strategiche.

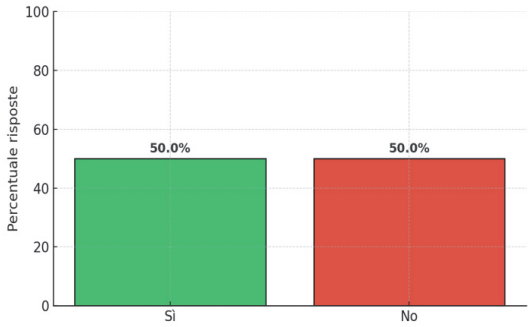
Le tre **aree più considerate** sono:

- **Automazione del servizio clienti.**
- **Sicurezza informatica.**
- **Ricerca e sviluppo.**

Questi ambiti riflettono un interesse combinato verso **efficienza operativa**, **difesa dei dati** e **spinta innovativa**.

. La sua azienda ha instaurato collaborazioni con università o centri di ricerca per lo sviluppo di soluzioni basate sull'Intelligenza Artificiale?

Le risposte si distribuiscono esattamente in modo paritario:

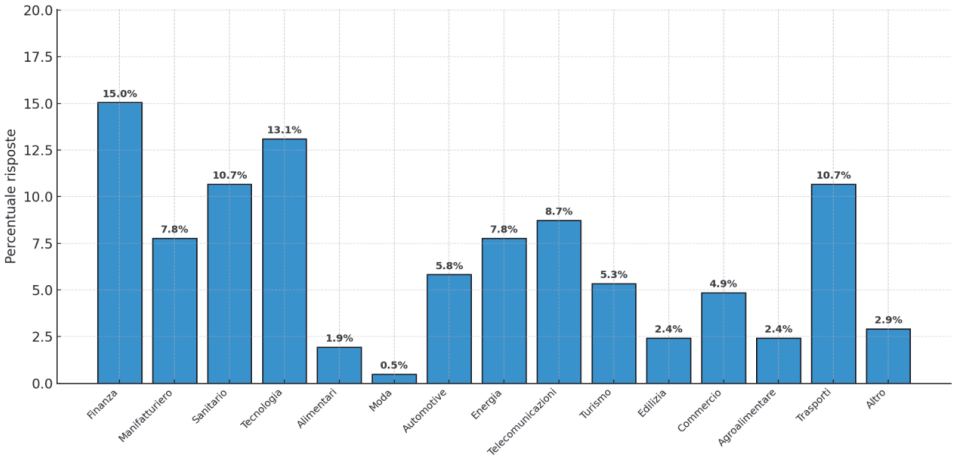


il 50% delle aziende (27 su 54) dichiara di aver attivato collaborazioni con università o centri di ricerca, mentre **l'altro 50% non ha intrapreso alcuna collaborazione.**

Questa perfetta simmetria evidenzia due scenari:

- da un lato, una metà del campione ha già colto le opportunità offerte dall'ecosistema della ricerca, verosimilmente per accedere a competenze avanzate, tecnologie emergenti o per beneficiare di finanziamenti e *partnership* pubblico-private;
- dall'altro, una quota equivalente rimane distaccata dal mondo della ricerca, forse per motivi culturali, di priorità aziendali o per assenza di connessioni strutturate.

. Qual è il settore industriale in cui ritiene che l'Intelligenza Artificiale possa avere il maggiore impatto attuale o potenziale futuro per la sua azienda? *(possibili più risposte)*

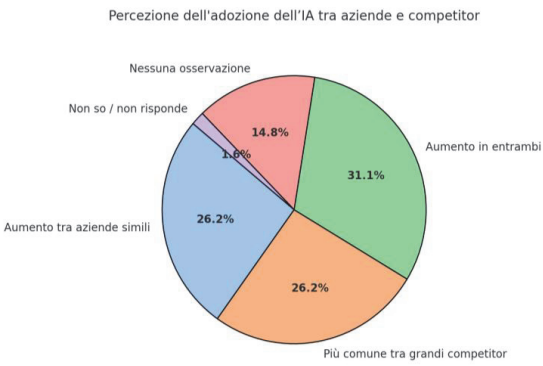


Le risposte a questa domanda indicano una percezione molto ampia e distribuita dell'impatto dell'IA nei vari settori economici, con una chiara predominanza di alcuni ambiti strategici.

Finanza e servizi bancari (31 risposte) è il settore più citato, probabilmente per la forte digitalizzazione già in atto e l'alto potenziale di automazione e data *analytics*.

Seguono **tecnologia ed elettronica** (27), **farmaceutico e sanitario** (22) e **trasporti e logistica** (22), tutti settori in cui l'IA può abilitare processi predittivi, ottimizzazione, diagnosi o efficienza operativa. Settori come **energia e ambiente**, **telecomunicazioni** e **manifatturiero** sono anch'essi ben rappresentati, confermando un impatto trasversale.

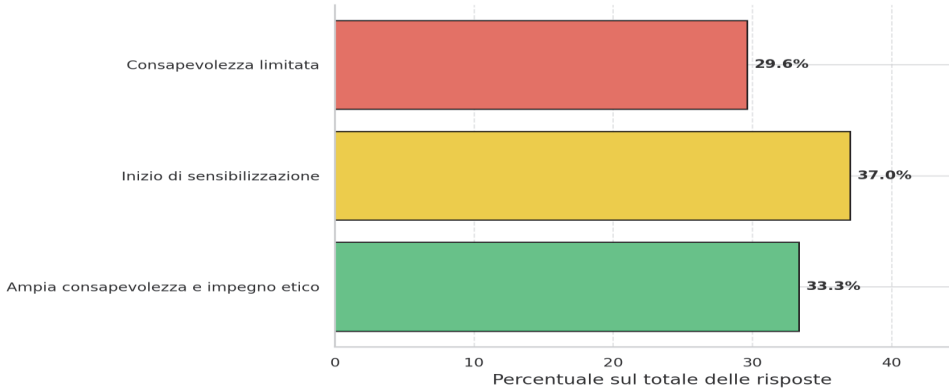
. Ha osservato un aumento dell'adozione dell'Intelligenza Artificiale tra le aziende simili alla sua o è più comune tra competitor più grandi?



La maggior parte dei rispondenti (31%) ha osservato una **crescita dell'adozione dell'IA sia tra aziende simili che tra grandi competitor**.

Questo dato indica una diffusione trasversale dell'innovazione, che sta toccando **sia realtà strutturate che imprese più agili** e suggerisce che **l'adozione dell'IA non è più un fenomeno elitario**, ma sta guadagnando trazione anche tra le **PMI**, seppur con differenze nei ritmi e nelle modalità.

. C'è una consapevolezza tra i suoi dipendenti riguardo alle opportunità e alle sfide etiche legate all'utilizzo dell'Intelligenza Artificiale?



Le risposte rivelano una distribuzione equilibrata tra i diversi livelli di consapevolezza etica:

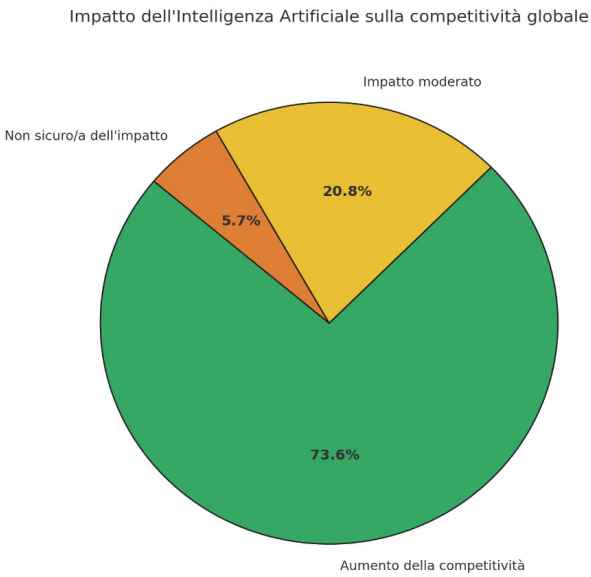
Il **33%** dei rispondenti riporta una *"ampia consapevolezza e impegno attivo"*, segnalando aziende già mature dal punto di vista della responsabilità digitale.

Il **37%** si trova nella fase iniziale di sensibilizzazione, indicando che **oltre due terzi del campione riconosce la necessità di affrontare le implicazioni etiche dell'IA**.

Tuttavia, circa un **30%** presenta una consapevolezza ancora limitata, suggerendo **un bisogno diffuso di formazione e cultura etica** all'interno delle organizzazioni.

Questi dati sottolineano l'importanza di accompagnare l'adozione dell'IA con iniziative di *awareness*, coinvolgimento e responsabilizzazione etica.

. Come ritiene che l'Intelligenza Artificiale possa influenzare la competitività della sua azienda a livello globale?

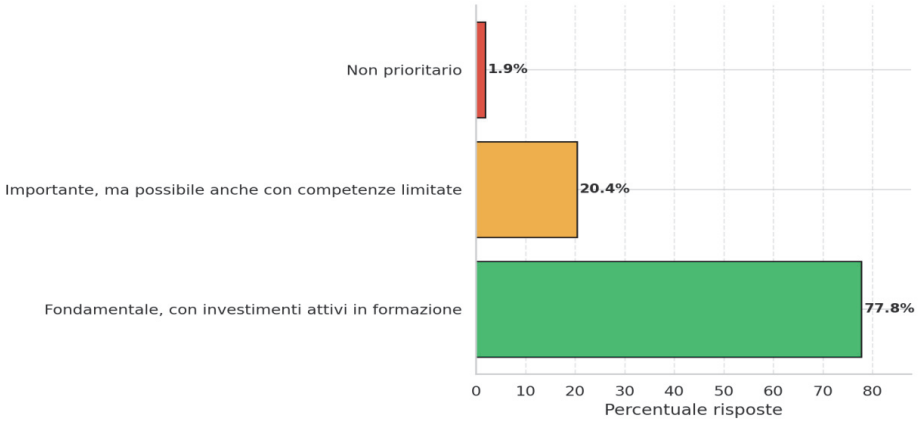


La stragrande maggioranza delle aziende (circa 74%) riconosce nell'Intelligenza Artificiale un **fattore di incremento della competitività globale**, soprattutto grazie a innovazione e maggiore efficienza operativa.

Un altro 21% segnala un impatto moderato, probabilmente in attesa di una maggiore maturità tecnologica o di condizioni di contesto più favorevoli.

Solo il 6% dei rispondenti esprime incertezza sull'effettivo impatto dell'IA, suggerendo che il valore strategico dell'intelligenza artificiale è ormai ampiamente riconosciuto nel panorama aziendale.

. Come valuta il ruolo delle competenze e della formazione nel facilitare l'adozione dell'Intelligenza Artificiale?



La formazione risulta essere un elemento **strategico** per l'adozione dell'Intelligenza Artificiale.

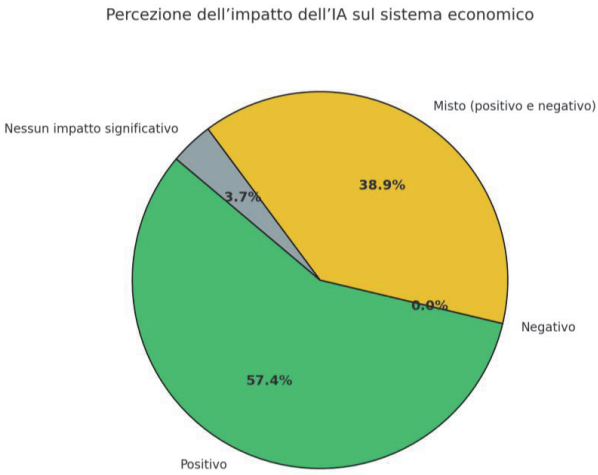
Ben **il 78%** delle aziende considera la formazione **fondamentale**, sostenendo attivamente i propri dipendenti con investimenti specifici.

Un altro **20%** riconosce comunque l'importanza delle competenze, pur ritenendo possibile l'adozione anche con capacità (skill) limitate, probabilmente ricorrendo a soluzioni esterne o semplificate.

Solo un'azienda su 54 (**2%**) non considera la formazione una priorità, un dato che conferma come la **capacità di aggiornamento delle risorse umane sia ritenuta cruciale** per restare competitivi nell'era dell'IA.

. Quale sarà, a suo parere, l'impatto dell'Intelligenza Artificiale sul sistema economico in generale?

La visione sull'IA come leva per il cambiamento economico è ampiamente positiva.

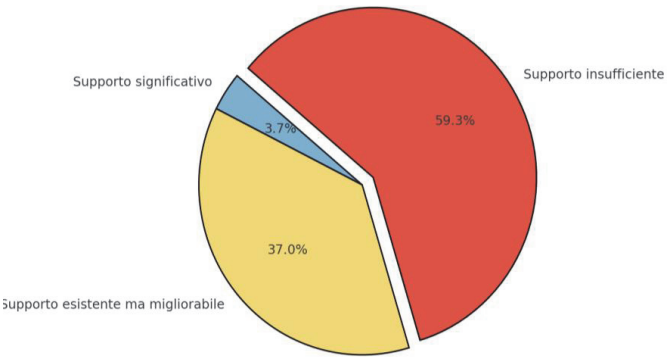


Il **57%** delle aziende intervistate ritiene che l'IA avrà un **impatto positivo sul sistema economico**.

Un altro **39%** adotta una prospettiva più bilanciata, riconoscendo **aspetti sia positivi che negativi** (es. efficienza vs. occupazione).

Il dato complessivo sottolinea un **alto livello di fiducia nel valore trasformativo dell'IA** a livello macroeconomico, pur con un certo grado di cautela in termini di effetti collaterali.

. Ritiene che ci sia un adeguato sostegno da parte del Governo per incoraggiare l'adozione dell'Intelligenza Artificiale nella sua industria?



I dati raccolti evidenziano una percezione critica del ruolo del governo nel promuovere l'adozione dell'IA:

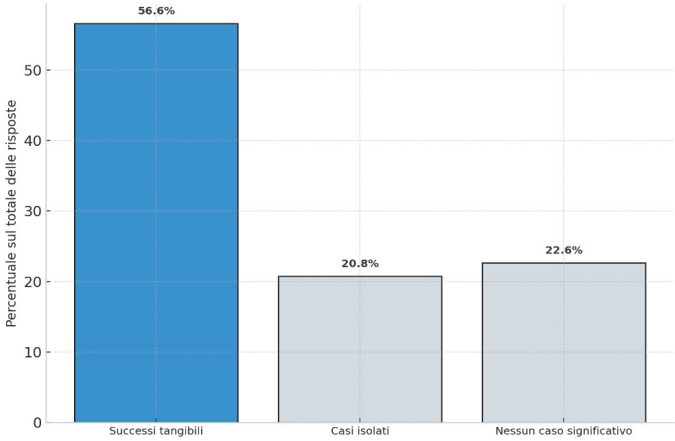
- **Solo il 4% delle aziende dichiara di percepire un sostegno istituzionale significativo**, tramite incentivi o strumenti finanziari mirati.

- **Una parte consistente (37%) riconosce l'esistenza di alcune iniziative**, ma ne sottolinea l'insufficienza o la **necessità di miglioramento.**

- **La maggioranza relativa (59%) ritiene che il sostegno governativo sia ancora insufficiente**, segnalando un forte bisogno di politiche pubbliche più strutturate e accessibili.

Il dato suggerisce l'urgenza di rafforzare le strategie nazionali sull'IA, anche per ridurre il divario competitivo a livello internazionale.

. Ha esperienze o conosce casi in cui l'Intelligenza Artificiale ha portato a miglioramenti significativi nelle operazioni o nei prodotti di aziende simili alla sua?

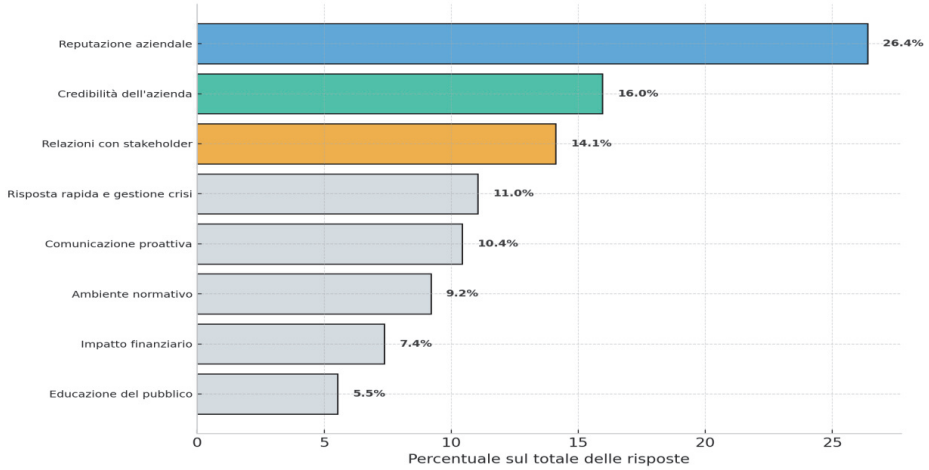


Le risposte evidenziano che oltre la metà delle aziende (57%) ha osservato casi concreti di successo nell'uso dell'IA. Ciò suggerisce un effetto positivo già misurabile in termini di:

- Efficienza operativa,
- Innovazione di prodotto,
- Ottimizzazione dei processi.

Il 20% ha riportato esperienze solo parziali o isolate, mentre il restante **22% non ha ancora riscontrato casi significativi**, segnalando che il grado di maturità dell'adozione varia ancora notevolmente tra le aziende.

. *Su quali ambiti si registrerà, a suo parere, l'impatto dell'Intelligenza Artificiale e della diffusione di* **fake news** *sulla gestione dell'informazione* **corporate?**



Le aziende riconoscono in larga misura l'impatto dell'IA e delle *fake news* su elementi cruciali della comunicazione corporate.

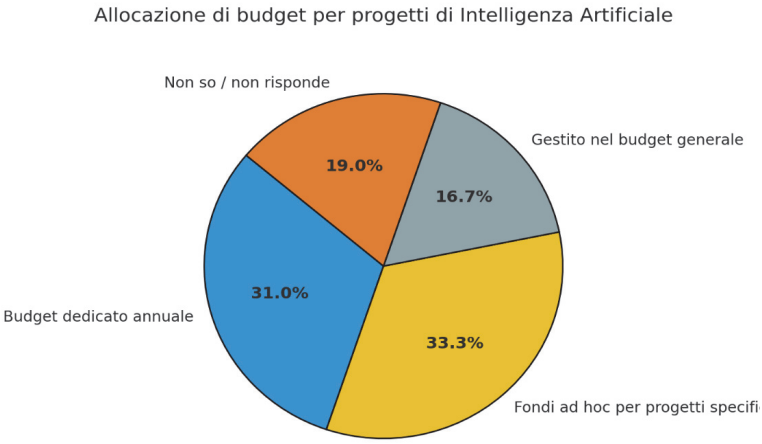
La **reputazione aziendale** (43 risposte) e la **credibilità** (26) emergono come gli ambiti più vulnerabili, indicando la centralità del trust nel contesto digitale.

Anche le **relazioni con gli stakeholder** (23), così come la capacità di **risposta in situazioni di crisi** (18), sono considerate aree sensibili.

Nel complesso, i dati sottolineano l'esigenza di **strategie aziendali strutturate per la gestione dell'informazione** in un ecosistema digitale reso complesso da IA e disinformazione.

SEZIONE 2: INVESTIMENTI E CONFORMITÀ ALLE NORMATIVE SULL'INTELLIGENZA ARTIFICIALE

. La sua azienda ha stanziato budget specifici per progetti di Intelligenza Artificiale?



L'analisi rivela che le imprese mostrano un **approccio variegato alla pianificazione finanziaria per l'IA**.

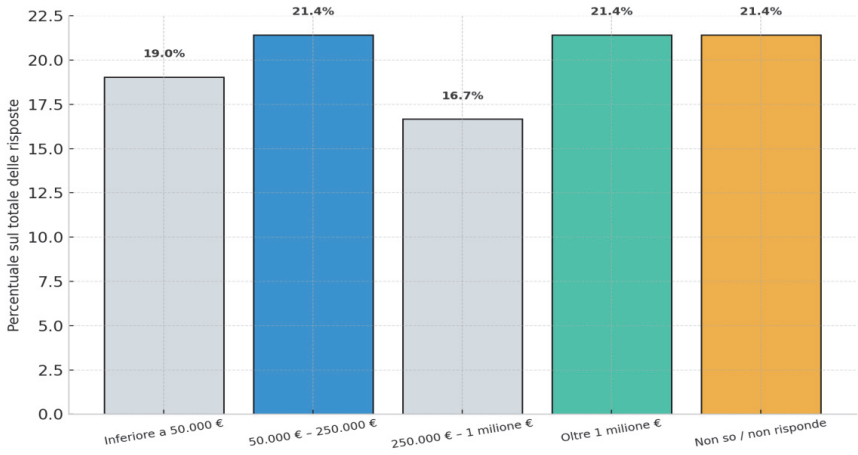
Il **31%** ha stanziato un **budget annuale dedicato**, segnalando una visione strutturata e strategica.

Una quota simile (**33%**) ricorre a **finanziamenti ad hoc per progetti specifici**, suggerendo flessibilità nella gestione.

Il **17%** integra gli investimenti in IA nel **budget generale**, senza una linea autonoma.

In sintesi, **oltre metà delle aziende che hanno adottato IA dispongono di forme specifiche di finanziamento**, un segnale importante di maturità organizzativa.

. Qual è l'ammontare degli investimenti realizzati dalla sua azienda in tecnologie e soluzioni basate sull'Intelligenza Artificiale nell'ultimo anno?



L'analisi delle risposte mostra una **distribuzione piuttosto equilibrata** degli investimenti effettuati dalle aziende in tecnologie basate sull'Intelligenza Artificiale:

- Il **21,4%** delle imprese ha dichiarato investimenti compresi tra **50.000 e 250.000 euro**.
- Un altro **21,4%** ha superato il milione di euro, dimostrando un impegno significativo e strutturato.
- Alla stessa percentuale (**21,4%**) appartiene anche il gruppo di aziende che **non ha fornito una stima** degli investimenti effettuati.

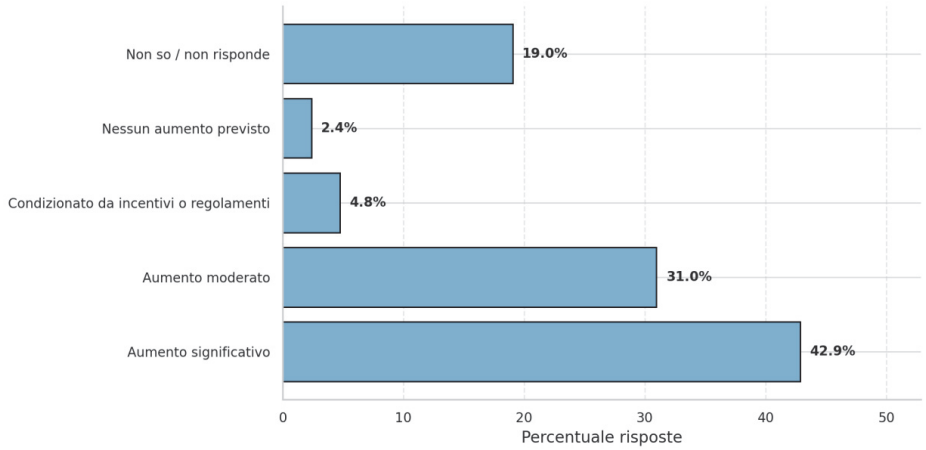
Seguono:

- Il **19%** con investimenti **inferiori a 50.000 euro**, che evidenziano un approccio ancora esplorativo o sperimentale.

- Solo il **16,7%** ha indicato una spesa tra **250.000 e 1 milione di euro**.

Questi dati confermano che **una parte consistente delle aziende è già attivamente impegnata in investimenti rilevanti**, anche se una quota non trascurabile non ha ancora piena visibilità o controllo sui fondi destinati all'IA.

. La sua azienda prevede di aumentare ulteriormente gli investimenti in Intelligenza Artificiale nel prossimo triennio, tenendo conto dell'evoluzione normativa?



Le intenzioni di investimento per il prossimo triennio segnalano una prospettiva di consolidamento e crescita:

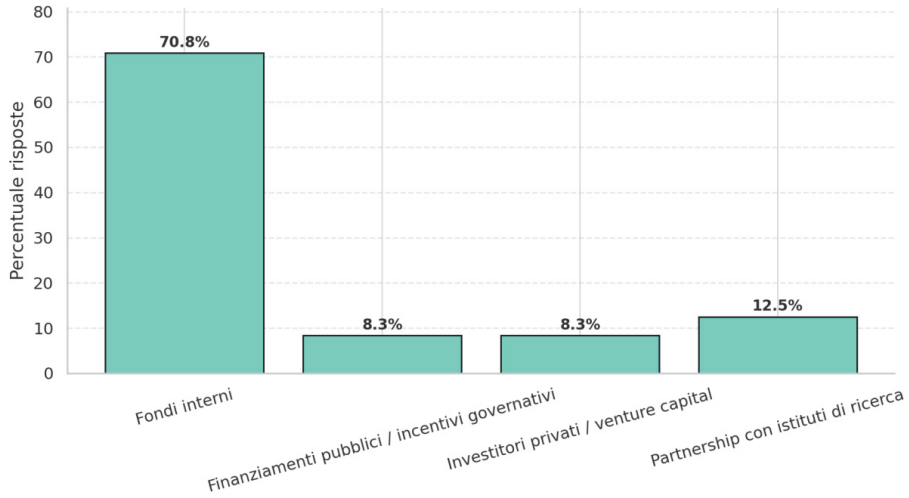
- **Circa il 43% prevede un aumento significativo degli investimenti**, mentre un ulteriore **31% indica un incremento moderato**.

- Solo il **2% esclude nuovi investimenti**, e il **5% li subordina a incentivi pubblici** o **sviluppi normativi**.

- **L'incertezza riguarda il 19% degli intervistati**, suggerendo un'area di attenzione in termini di pianificazione strategica o conoscenza delle opportunità.

In sintesi, oltre il 70% delle aziende manifesta una chiara volontà di intensificare le risorse dedicate all'IA, consolidando la transizione verso una trasformazione *data-driven*.

. Quali fonti di finanziamento ha utilizzato la sua azienda per sostenere gli investimenti in Intelligenza Artificiale?

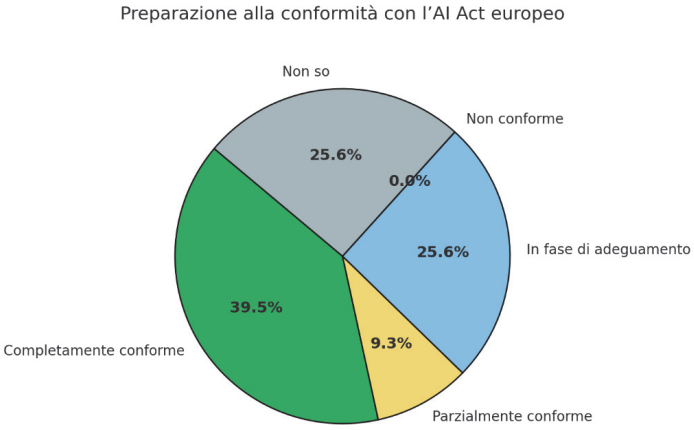


Il finanziamento dell'IA proviene principalmente da **risorse interne**, con scarsa incidenza di capitali esterni:

- Circa il **71%** delle aziende ha finanziato i progetti **con fondi propri**, confermando un forte impegno autonomo.
- Solo **l'8,3%** ha attinto a **finanziamenti pubblici** o **venture capital**, a testimonianza di una **limitata capitalizzazione di opportunità esterne**.
- Il **12,5%** ha stabilito **partnership con centri di ricerca**, segnalando un modello collaborativo in crescita, seppur minoritario.

Questa situazione evidenzia un potenziale non ancora sfruttato in termini di bandi, incentivi e alleanze strategiche, soprattutto per le imprese di piccole e medie dimensioni.

. In che misura la sua azienda è preparata per conformarsi all'AI Act europeo?

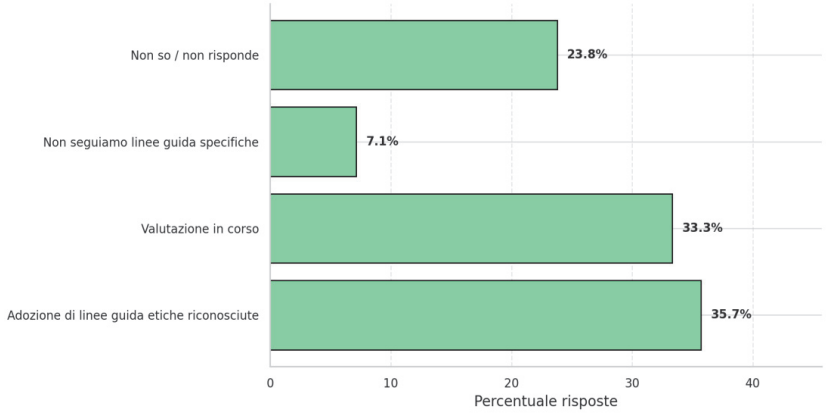


La preparazione normativa all'AI Act è **ancora in fase evolutiva**.

- Il **39,5%** delle aziende si dichiara **completamente conforme**, mentre un altro **9%** è **parzialmente conforme**.
- Il **25%** è in **fase di adeguamento**, indicando attenzione ma non ancora piena conformità.
- Il **25%** dei rispondenti **non sa valutare il proprio livello di conformità**, suggerendo **una necessità di maggiore informazione e assistenza tecnica**.
- **Nessuna azienda ha dichiarato di essere esplicitamente "non conforme"**.

Questa analisi suggerisce che, sebbene vi sia un'attenzione crescente al quadro normativo, serve un rafforzamento delle competenze giuridiche e procedurali per garantire una piena aderenza al nuovo regolamento europeo.

. La sua azienda segue linee guida etiche o standard internazionali per l'utilizzo dell'Intelligenza Artificiale (es. IEEE, OECD, etc.)?

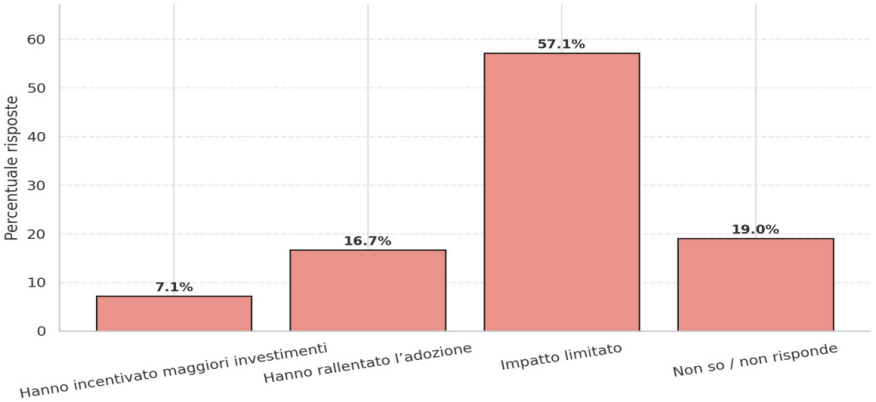


Le aziende stanno gradualmente integrando principi etici nella *governance* dell'IA.

- Il **36%** dichiara di seguire linee guida riconosciute a livello internazionale (es. OECD, IEEE, etc.), un segnale positivo di responsabilizzazione.
- Un ulteriore **33%** è in fase di valutazione, il che lascia intravedere una futura estensione dell'adozione.
- Solo il **7%** non segue alcuna linea guida specifica, mentre il **24%** non ha una posizione chiara o non è informato.

Il dato complessivo mostra una crescente attenzione ai temi etici e normativi, anche se restano spazi da colmare per una vera convergenza su standard globali.

. In che modo le normative attuali (nazionali/internazionali) sull'Intelligenza Artificiale hanno influenzato le strategie e gli investimenti della sua azienda?



Le **normative sull'Intelligenza Artificiale** esercitano un **influsso moderato** e **differenziato** sulle strategie aziendali.

- **Per oltre il 57% delle aziende, le normative hanno avuto un impatto limitato**, indicando che l'adozione dell'IA è trainata da motivazioni autonome più che da vincoli esterni.
- **Quasi il 17% segnala che la regolamentazione ha rallentato l'adozione**, probabilmente per la complessità o i costi associati alla *compliance*.
- **Solo il 7% dichiara che le normative hanno stimolato maggiori investimenti**, spesso per conformarsi e acquisire vantaggi competitivi.
- **Il 19% non sa valutare l'impatto**, riflettendo un gap informativo o strategico che potrebbe limitare la reattività normativa delle aziende.

CONCLUSIONI

L'anno 2025 segna un punto di svolta nel processo di maturazione dell'intelligenza artificiale a livello globale. L'IA si conferma non soltanto come tecnologia abilitante, ma come infrastruttura strategica, capace di influenzare dinamiche industriali, economiche, sociali e geopolitiche. In tale quadro, l'Unione Europea ha ribadito l'ambizione di definire una **"via europea" all'IA**, fondata su diritti fondamentali, sostenibilità, pluralismo e sicurezza. Un approccio che, per essere efficace, deve però tradursi in capacità produttiva, interoperabilità, investimento e formazione.

Il **modello europeo**, fondato sull'etica e sulla regolazione, rischia infatti di rimanere incompleto senza una **massa critica tecnologica e industriale**. A fronte di giganti privati (Stati Uniti) e apparati statali centralizzati (Cina), l'Europa sconta una frammentazione operativa, pur avendo positivamente accelerato su strumenti come l'**AI Act**, i fondi **InvestAI** e il piano strategico **AI Continent Plan**.

In questo contesto, l'Italia può contribuire con un'identità propria, valorizzando il suo potenziale **di integrazione verticale dell'IA in ambiti creativi, culturali e manifatturieri**, dove eccelle per *design*, artigianalità e capacità di personalizzazione.

La combinazione di creatività, ingegnosità e qualità estetico-funzionale offre uno spazio distintivo per un utilizzo originale dell'IA nei settori del *made in Italy*, della sanità personalizzata, del turismo, dell'agroalimentare e del patrimonio culturale.

La **diffusione dell'IA in Italia** continua, tuttavia, a risentire della struttura economica del proprio tessuto industriale, costituito per lo più da PMI che spesso possono avere difficoltà nel reperire risorse e competenze e che devono creare una loro propria cultura

digitale necessarie per un'adozione piena. Per superare questa asimmetria e consentire, all'interno del contesto industriale italiano, lo sviluppo pieno di questa tecnologia e di tutto il potenziale che porta con sé, occorre favorire politiche pubbliche che incentivino l'adozione reale delle tecnologie nei processi produttivi, rafforzare i competence center e promuovere modelli di collaborazione interaziendale in logica di filiera.

Dal punto di vista **tecnologico**, il 2025 ha consolidato alcune direttrici: l'**IA generativa** e l'**IA collaborativa** si sono diffuse nei settori dei servizi, della manifattura e della pubblica amministrazione, abilitando nuove forme di interazione uomo-macchina. Crescono anche le applicazioni in ambito **cybersecurity**, **identità digitale** e **manutenzione predittiva**, ma rimangono aperti i nodi relativi alla sostenibilità energetica e alla trasparenza algoritmica.

Dal punto di vista **etico-sociale**, si rafforza la consapevolezza che l'IA non può essere lasciata esclusivamente nelle mani dei produttori di tecnologia: è urgente costruire una ***governance* multilivello**, che integri visioni giuridiche, filosofiche e democratiche.

L'adozione massiva di IA comporta rischi concreti di disuguaglianza, discriminazione e perdita di *agency* umana. Serve dunque una nuova cultura dell'interazione uomo-macchina, fondata su formazione, vigilanza e pluralismo.

Il **mercato del lavoro continua a trasformarsi**: emergono nuove professioni ibride, mentre il *mismatch* tra competenze richieste e disponibili rischia di aggravare le disuguaglianze. Per questo, il **capitale umano** va posto al centro delle strategie: la formazione tecnica, la diffusione della cultura digitale e l'alfabetizzazione sull'uso dell'IA devono essere priorità per il Paese.

Infine, emerge con forza il bisogno di **visione strategica e convergenza internazionale**. In assenza di una *governance* condivisa - soprattutto tra le democrazie liberali - il rischio è un mondo frammentato, polarizzato tra modelli divergenti. In questa prospettiva, la proposta di un **centro di ricerca europeo sull'IA** - **un "CERN dell'intelligenza artificiale"** - torna di attualità: esso potrebbe catalizzare competenze, risorse, standard e policy comuni, diventando fulcro della via europea all'intelligenza artificiale.

In sintesi:

- L'IA è ormai una forte **leva trasformativa e di competitività**.
- L'**Europa** ha bisogno di potenziare e sviluppare appieno la propria **capacità produttiva e infrastrutturale**, oltre che regolatoria.
- L'**Italia** può emergere come *hub* di **integrazione creativa e settoriale dell'IA**, con applicazioni distintive in settori ad alto valore simbolico ed economico.
- Servono **politiche di lungo periodo** che favoriscano l'adozione dell'IA e la rendano accessibile anche alle PMI e sostenibile per l'ambiente.
- Una *governance* dell'IA **etica, inclusiva e trasparente** è la condizione per uno sviluppo tecnologico realmente democratico e sostenibile, che ponga sempre l'uomo al centro.

RIFERIMENTI BIBLIOGRAFICI

A Declaration for the Future of the Internet, April 28, 2022.
https://www.state.gov/declaration-for-the-future-of-the-internet

STANFORD HAI, *Artificial Intelligence Index Report 2025*.
https://hai.stanford.edu/ai-index/2025-ai-index-report

AMATO, G. & CONTUCCI, P. (2025), "La ricerca in intelligenza artificiale tra libertà e potere", *Rivista Il Mulino online*, 12 marzo 2025.

ALMADA M., RADU A. (2024), "The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy", *German Law Journal*, 2024, pp. 1 e ss.

BONADIO E., MCDONAGH L. (2020), "Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity", *Intellectual Property Quarterly*, 2020, pp. 112-137.

CARTA M. (2024), "Il Regolamento UE sull'Intelligenza Artificiale: alcune questioni aperte", *Eurojus.it*, 2, 2024, in part. p. 190.

CHIRIATTI, M., GANAPINI, M., PANAI, E., UBIALI, M., RIVA, G. (2024), "The case for human–AI interaction as system 0 thinking", *Nature Human Behaviour*, 8(10), 1829-1830.

CONTUCCI, P. (2024), "Intelligenza Artificiale: rischio energetico e climatico", *Rivista Il Mulino online*, 2 luglio 2024.

DELL'ACQUA, F., AYOUBI, C., LIFSHITZ -ASSAF, H., SADUN, R., MOLLICK, E. R., MOLLICK, L., HAN, Y., GOLDMAN, J., NAIR, H., TAUB, S., & LAKHANI, K. R. (2025), "The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise", *Harvard Business School Working Paper* No. 25-043.

D'ONOFRIO, M. (2024), "Azioni e creazioni dell'intelligenza artificiale", *Tecnologie e diritto*, 1/2024.

DI BENEDETTO, C. (2022), "Creatività artificiale e diritti d'autore", *AIDA - Annali italiani del diritto d'autore, della cultura e dello spettacolo*, Vol. 31/1, 2022, pp.388-402.

FERRETTI V. F. (2022), "Intelligenza artificiale e diritto d'autore: quale tutela per il robot creatore?", *Rassegna del diritto della moda e delle arti*, 2022, pp. 68-106.

FINOCCHIARO G. (2023), *Diritto di internet*, Zanichelli Torino, IV ed., 2023.

FROSIO G. (2020), "L'(I)Autore inesistente: una tesi tecno-giuridica contro la tutela dell'opera generata dall'intelligenza artificiale", *AIDA - Annali italiani del diritto d'autore, della cultura e dello spettacolo*, Vol. 29, 2020, pp. 52-91.

GARG A., KITSARA I., BÉRUBÉ S. (2025), "The hidden cost of AI: Unpacking its energy and water footprint", *OECD.AI Policy Observatory*, 26 February 2025.

GITTO P. (2024), "*New York Times* v. OPENAI, Microsoft *et al*.: conflitti attuali fra intelligenza artificiale e diritto d'autore", *Giustizia Civile*, 2024, pp. 185 ss.

GUIZZARDI S. (2018), "La protezione d'autore dell'opera dell'ingegno creata dall'Intelligenza Artificiale", *AIDA - Annali italiani del diritto d'autore, della cultura e dello spettacolo*, Vol.28, 2018, pp. 1-31.

INTERNATIONAL HOSPITAL FEDERATION, "Unlocking the potential of AI: Emerging opportunities, challenges, risks and insights for healthcare leaders (YEL 2024)", (by M. Abdelhakim (Egypt), S. Addessi (USA), M. Padi Amoatey (Ghana), A. Girouard (Canada), J. Leido (USA) and A. Verissimo (Brazil)), 19 September 2024.

LANZALONGA, F., MARSEGLIA, R., IRACE, A., & BIANCONE, P. P. (2024), "The application of artificial intelligence in waste management: understanding the potential of data-driven approaches for the circular economy paradigm", *Management Decision*, February.

MARSEGLIA, G.R., REALI, A., PREVITALI, P. (2025), *Socio-economic Impact of Artificial Intelligence: A European Management Perspective*, Springer Verlag.

The New York Times (2025), "Saying 'Thank You ' to ChatGPT Is Costly. But Maybe It's Worth the Price", 24 April 2025.

NIZZA, U. (2024), "Assessing the Impact of the European AI Act on Innovation Dynamics: Insights from Artificial Intelligences", Northwestern University 2024.

NOVELLI C. *et al*. (2024), "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity", *Computer Law & Security Review*, 55, 2024.

OECD (2024), *Explanatory memorandum on the updated OECD definition of an AI system*, Policy Paper OECD, Artificial Intelligence Papers, 5 March 2024.

SCHNEIDER, E. (2024), "IA Act e i sistemi di rischio: lungimiranza o nostalgia?", Osservatorio sull'AI Act, *IRPA*, 12 dicembre 2024.

SCHREPEL T. (2025), "Decoding the AI Act: Implications for Competition Law and Market Dynamics", *Journal of Competition Law & Economics*, 4 March 2025.

SIEGMANN C., ANDERLJUNG M. (2022), *The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market*, Centre for the Governance of the AI, August 2022.

SPANÒ, R. (2016), *L'evoluzione dei sistemi di Management Accounting nelle aziende sanitarie - Accountability e fattori di complessità*, G. Giappichelli Editore Torino, 2016.

WORLD ECONOMIC FORUM (2025), *Artificial Intelligence's Energy Paradox: Balancing Challenges and Opportunities*, White Paper, January 2025.

Osservatorio Permanente

sull'Adozione e l'Integrazione della Intelligenza Artificiale

(IA²)

Rapporto Intelligenza Artificiale

2025

COMPONENTI DEL COMITATO SCIENTIFICO

Michele Abrusci, *Senior Professor*, Logica e Filosofia della Scienza, Università degli Studi Roma Tre.

Giuliano Amato, *Presidente del Comitato Scientifico*; Presidente Onorario, Aspen Institute Italia; Presidente Emerito, Corte costituzionale.

Angelo Federico Arcelli, Professore Straordinario di Economia delle Istituzioni Finanziarie e Internazionali, Università degli Studi Guglielmo Marconi; *Senior Fellow*, CIGI Center for International Governance Innovation.

Alessandro Armando, Ordinario, Dipartimento di Informatica, Bioingegneria, Robotica e Ingegneria dei Sistemi, Università degli Studi di Genova.

Paolo Benanti, Professore Straordinario di Etica della Tecnologia, Pontificia Università Gregoriana.

Patrizio Bianchi, Professore Emerito di Economia; Cattedra Unesco "Educazione, Crescita e Uguaglianza", Università degli Studi di Ferrara.

Gian Carlo Blangiardo, Professore Emerito di Demografia, Università degli Studi di Milano-Bicocca; già Presidente ISTAT.

Barbara Caputo, Ordinaria di Intelligenza Artificiale e Direttrice Hub AI@PoliTO, Politecnico di Torino.

Sabino Cassese, Giudice Emerito, Corte costituzionale; Professore Emerito, Scuola Normale Superiore di Pisa.

Pierluigi Contucci, Ordinario di Fisica Matematica, *Alma Mater Studiorum*, Università degli Studi di Bologna.

Anna Corrado, Magistrato Amministrativo, Tribunale Amministrativo Regionale per la Lombardia.

Carmela Decaro Bonella, già Professoressa di Diritto Pubblico e Comparato, Luiss Guido Carli.

Giusella Finocchiaro, Ordinaria di Diritto Privato e di Diritto di Internet, *Alma Mater Studiorum*, Università degli Studi di Bologna.

Anna Gatti, Direttore, LIFT Lab - SDA Bocconi School of Management; Co-founder, Angel Investor, Non-Executive Director.

Michel Ghins, Professore Emerito di Filosofia della Scienza, Université Catholique de Louvain.

Luigi Gubitosi, Docente di Finanza Aziendale Internazionale, Luiss Guido Carli.

Massimo Massella Ducci Teri, Avvocato Generale dello Stato Emerito.

Franco Massi, Magistrato; Segretario Generale, Corte dei conti; Presidente OIV, Ministero della Difesa.

Marco Mayer, Docente Master *CyberSecurity*, Luiss Guido Carli.

Giancarlo Montedoro, Presidente di Sezione, Consiglio di Stato; già Consigliere Giuridico del Presidente della Repubblica.

Jacques Moscianese, Executive Director Institutional Affairs Intesa Sanpaolo.

Carlo Nardello, Professore di *Marketing* Digitale, Sapienza Università di Roma.

Alessandro Pajno, Docente di Diritto Processuale Amministrativo, LUISS Guido Carli; Presidente Emerito, Consiglio di Stato.

Giulio Perani, Dirigente di Ricerca, ISTAT.

Angelo Maria Petroni, Segretario Generale, Aspen Institute Italia; Ordinario di Logica e Filosofia della Scienza, Sapienza Università di Roma.

Andrea Ripa, Vescovo Titolare di Cerveteri; Segretario del Supremo Tribunale della Segnatura Apostolica.

Mariarosaria Taddeo, Professore di *Digital Ethics* e *Defence Technologies*, Università di Oxford.

COMPONENTI DEL COMITATO TECNICO

Matteo Boaglio, Head of Institutional Special Projects and Policies Intesa Sanpaolo.

Silvia Castagna, Responsabile Relazioni Istituzionali e Grandi Clienti, DOXA Istituto Ricerche Statistiche e Analisi Opinione Pubblica.

Valerio Cencig, Executive Director Compliance Digital Transformation Intesa Sanpaolo.

Marco Ditta, Executive Director Data & Artificial Intelligence Office Intesa Sanpaolo.

Lorenzo Iannarilli, Cultore della Materia in *Marketing and Digital Communication*, Università LUMSA.

Laura Li Puma, Senior Director Applied Research and Innovation Hubs Intesa Sanpaolo Innovation Center.

Giuseppe Roberto Marseglia, *Aspen Junior Fellows Alumnus*, CEO, Daat Consulting.

Per Aspen Institute Italia, Roberto Billiani e Monica Coppi.

L'**Osservatorio Permanente sull'Adozione e l'Integrazione della Intelligenza Artificiale (IA²)**, fondato nel 2023 e presieduto da Giuliano Amato, costituisce l'istituzionalizzazione di una serie di attività di Aspen Institute Italia sull'Intelligenza Artificiale, tra le quali le due conferenze internazionali "Ethics and Artificial Intelligence", tenutesi a Venezia nel 2021 e 2022, alle quali il Presidente Mattarella ha voluto conferire il suo Alto Patronato (2021) e la Medaglia del Presidente della Repubblica (2022). L'Osservatorio, sostenuto da Intesa Sanpaolo, è luogo di riflessione e di proposta nella tradizione di Aspen, caratterizzata dal libero confronto delle idee per il raggiungimento del bene comune, in una dimensione sia nazionale sia europea sia globale.

Aspen Institute Italia, fondato nel 1984, è un'associazione privata, indipendente, internazionale, apartitica e senza fini di lucro. La sua attività è caratterizzata dalla discussione e dall'approfondimento di temi strategici per la società contemporanea nonché dallo scambio di conoscenze, informazioni e valori.
Dalla sua fondazione, l'Istituto ha contribuito alla vita politica, economica e culturale italiana attraverso dibattiti, tavole rotonde, conferenze e seminari, alla sua rivista *Aspenia* (fondata nel 1995) e ad *Aspenia online* (fondata nel 2008) – così come mediante molteplici studi e progetti di ricerca.