



Digital

Makers & Shapers

AI GENERATIVA: LA RICERCA DI UNA REGOLAMENTAZIONE E DI UNA LEADERSHIP INDUSTRIALE IN EUROPA



INDICE

SINTESI	4
INTRODUZIONE	6
COS'È E COME FUNZIONA L'AI GENERATIVA	9
DINAMICHE SOCIOECONOMICHE DELL'AI GENERATIVA	12
SVILUPPI DEL MERCATO	12
PROBLEMI DI CONCORRENZA	15
IMPATTI ECONOMICI	17
DAI RISCHI ALLA REGOLAMENTAZIONE E ALLE PROSPETTIVE DELL'UE	26
RISCHI POSTI DALL'AI GENERATIVA	26
LA REGOLAMENTAZIONE: UN'ANALISI GLOBALE SELETTIVA	28
REGOLAMENTAZIONE: L'AI ACT DELL'UE	32
PROSPETTIVE DELL'UE SULL'AI GENERATIVA	37
CONSIDERAZIONI COMPARATIVE CONCLUSIVE	40
TECNOLOGIA, DINAMICHE SOCIOECONOMICHE E REGOLAMENTAZIONE: GLI SCENARI	42
PRINCIPALI AMBITI DI INCERTEZZA	42
GLI SCENARI PROPOSTI	44
TRAME DEGLI SCENARI	45
VALUTAZIONE DEGLI SCENARI E CONCLUSIONI	49
RINGRAZIAMENTI	55
BIBLIOGRAFIA/SITOGRAFIA	56

SINTESI

Il futuro dell'AI generativa in Europa presenta uno scenario caratterizzato da un potenziale di miglioramento, sfide normative e dalla ricerca di una leadership di settore. Mentre l'Europa esplora le dinamiche in evoluzione dell'AI generativa, dal presente report emergono diversi fattori chiave.



L'AI generativa annovera modelli che sintetizzano nuovi audio, codici, immagini, testi, video, progetti, materiali e altre rappresentazioni strutturali provenienti da grandi volumi di dati spesso non strutturati e non etichettati per mezzo di grandi reti neurali addestrate con algoritmi di machine learning.



Lo sviluppo dell'AI generativa ha avuto una repentina accelerazione dopo il lancio di ChatGPT da parte di OpenAI nel novembre del 2022 con un ritmo di crescita molto sostenuto. Il ciclo di vita, la frequenza di utilizzo e la rapida integrazione nelle applicazioni software esistenti sono notevoli. Le soluzioni di AI generativa si stanno diffondendo e si evidenzia l'arrivo di nuovi player accanto ai giganti della tecnologia e un netto aumento degli investimenti lungo la catena del valore dell'AI generativa.



L'AI generativa promette di aumentare la produttività e di migliorare la qualità degli output e, potenzialmente, quella del lavoro in diverse mansioni. Sebbene permangano incertezze sulla dislocazione dei posti di lavoro e sulle implicazioni sociali, le prime indicazioni suggeriscono impatti positivi sulla crescita economica e sulle funzioni lavorative.



Attualmente l'Europa è in ritardo, rispetto alla concorrenza mondiale, nel campo dell'AI generativa con una significativa disparità nell'assegnazione di finanziamenti alle start-up europee rispetto alle controparti statunitensi e cinesi. L'assenza di un'iniziativa unitaria riguardo l'AI a livello europeo pone delle sfide all'Europa per affermarsi con prodotti e infrastrutture di AI "Made in Europe".



Non è in gioco solo l'economia. Come abbiamo discusso nelle sezioni precedenti, si prevede che l'AI generativa avrà un impatto importante sulla società e un'area a rischio è la cultura. Il dominio assoluto delle grandi aziende statunitensi e cinesi di AI è sconcertante. Non disponendo di piattaforme di AI europee valide, gli utenti europei dovranno accettare le imposizioni etico-culturali dell'AI sviluppata in altre parti del mondo e la diversità linguistico-culturale europea potrebbe risentirne.

A fronte di questa disparità, si possono trarre dal report le principali implicazioni seguenti:



Equilibrio tra regolamentazione e innovazione. I responsabili politici sono invitati ad adottare misure proattive per trovare un punto di equilibrio tra l'innovazione e le considerazioni etiche. Piattaforme di collaborazione come l'AI Safety Summit 2023 facilitano il dialogo internazionale sulla mitigazione dei rischi e la massimizzazione delle opportunità in relazione alle tecnologie di AI. Se da un lato la regolamentazione è fondamentale per mitigare i rischi, dall'altro un suo eccesso può soffocare l'innovazione, soprattutto per le start-up più piccole. Suggerisce di realizzare sandbox e collaudi regolamentari per i ricercatori e le aziende europee di AI generativa al fine di sperimentare in modo responsabile in conformità alla legge europea sull'AI.



Finanziamenti. L'aumento dei finanziamenti è essenziale, ma dovrebbe essere mirato a ecosistemi aperti e alla collaborazione, non solo a singoli grandi player tecnologici. Iniziative come quella proposta di GenAI4EU, per creare centri intersettoriali, possono contribuire a rendere più democratico l'accesso alle risorse. I progetti rivelati dalla Commissione Europea, il 24 gennaio 2024, per supportare le start-up e le PMI di Intelligenza Artificiale sono un primo passo, ma c'è ancora molto da fare e i finanziamenti concessi non sono sufficienti.



Problemi di concorrenza. Oltre alle iniziative volte alla regolamentazione come l'AI Act, è necessario affrontare le incertezze relative alle politiche di concorrenza e alle iniziative antitrust nel quadro normativo dell'AI generativa. Il rischio di pratiche di concorrenza sleale da parte degli operatori storici nel mercato dell'AI generativa, come l'accorpamento di prodotti o la stipula di accordi in esclusiva, deve essere mitigato. Le potenziali barriere all'accesso per i nuovi operatori devono essere abbassate favorendo l'accesso al talento e alla potenza di calcolo.

In conclusione, il futuro dell'AI generativa in Europa dipende dallo studio delle complessità normative, dalla promozione dell'innovazione e dagli sforzi di collaborazione per garantire uno sviluppo e una diffusione responsabili delle tecnologie di AI. Accogliendo queste sfide in modo proattivo, l'Europa può posizionarsi come leader nel plasmare il futuro dell'AI generativa, sostenendo al contempo i principi di affidabilità, trasparenza e benessere della società.

INTRODUZIONE

Alcuni sostengono che da novembre 2022, quando OpenAI ha rilasciato la prima versione pubblica di ChatGPT basata sui Large Language Models (di seguito semplicemente LLM), sia nata una nuova fase della rivoluzione dell'AI. ChatGPT ha raggiunto cento milioni di utenti in circa due mesi, diventando l'applicazione a più rapida crescita nella storia del Web. Dopo ChatGPT, sono seguite Sora e molte altre soluzioni simili, come Bard, Claude, Midjourney e vari altri strumenti per la creazione di contenuti.

ChatGPT e soluzioni simili è un tipo di AI generativa per scopi generali, alla quale, per semplicità, d'ora in poi ci riferiremo solo con AI generativa o GAI. Siamo consapevoli che l'AI generativa per scopi generali dovrebbe essere distinta dall'AI generativa per scopi speciali e anche che nel dibattito politico e nella letteratura scientifica vengono utilizzate in modo intercambiabile diverse espressioni, come "AI per scopi generali", "modelli di fondazione", "modelli di frontiera", "modelli linguistici di grandi dimensioni" o "Large Language Models". In questo report, per motivi di coerenza e brevità, utilizziamo l'espressione AI generativa (quasi sempre abbreviata in GAI) per riferirci in generale a modelli che sintetizzano nuovi audio, codici, immagini, testi, video, progetti, materiali e altre rappresentazioni strutturali a partire da grandi volumi di dati spesso non strutturati e non etichettati, sfruttando grandi reti neurali addestrate con algoritmi di machine learning. In questa definizione, la GAI è un insieme specifico di sistemi machine-based in grado di generare contenuti a partire da dati di input secondo un obiettivo.

I dati presentati successivamente in questo report sugli sviluppi del mercato mostrano una crescente diffusione delle soluzioni GAI, l'ingresso di nuovi player accanto ai giganti del settore tecnologico e un netto aumento degli investimenti lungo tutta la catena di valore GAI. Dopo essere stati addestrati, i modelli GAI mostrano la capacità

di generare contenuti che assomigliano molto ai modelli e alle informazioni presenti nei dati di addestramento. Le varianti basate sulla chat di questi modelli sono generalmente utilizzate anche per rispondere alle richieste degli utenti. Ad esempio, ChatGPT può generare testi e contributi di dialogo simili a quelli umani e Midjourney può generare immagini realistiche. Mentre i sistemi di intelligenza artificiale precedenti erano in grado di generare piccole quantità di contenuti (ad esempio, suggerendo modifiche ortografiche o stilistiche a un testo esistente o apportando modifiche alle immagini), la GAI è in grado di generare contenuti di alta qualità da zero, a partire da minime richieste¹. Tuttavia, come accade spesso quando le nuove tecnologie sembrano decollare, la domanda tipica che gli analisti si pongono è la seguente: siamo di fronte a un nuovo fervido interesse tecnologico destinato ad affievolirsi o si tratta di una vera e propria svolta che trasformerà radicalmente molti aspetti della vita economica e sociale?

A giudicare dalle prime analisi e dai primi studi disponibili (compresi quelli sperimentali), sembra che la GAI possa produrre un aumento sostanziale della produttività e un miglioramento della qualità dell'output e del lavoro, almeno per alcune tipologie di mansioni. Alcuni studi hanno rilevato che i LLM basati sulla GAI potrebbero avere un impatto sull'80% dei lavoratori negli Stati Uniti² e potrebbero aumentare il PIL globale annuale del 7% in dieci anni³. Uno studio incentrato sull'Italia ha stimato che l'AI generativa potrebbe aumentare la produttività per un importo pari al 18% del PIL o, in altre parole, potrebbe supplire fino a 5,7 miliardi di ore di lavoro⁴. Inoltre, anche se il dibattito accademico e le prove sono ancora in fase iniziale e i risultati non possono essere considerati definitivi, ci sono indicazioni che la GAI potrebbe non solo dislocare posti di lavoro, ma, in realtà, migliorare il lavoro degli esseri umani, compreso quello dei lavoratori meno qualificati, cosa che non è avvenuta con le precedenti ondate di automazione.

La GAI potrebbe avere enormi promesse, ma è ancora lontana dalla perfezione priva di potenziali rischi. In primo luogo, c'è il fenomeno ben documentato degli errori negli output dei modelli GAI⁵, definiti "allucinazioni". I rischi per la società includono l'aumento della discriminazione, della disinformazione e dell'eccessiva fiducia nei contenuti della GAI⁶. Solo per citare un esempio, la provenienza dei contenuti è fonte di preoccupazione. Gli esperimenti hanno dimostrato che i consumatori umani non sono in grado di individuare se un contenuto è stato generato dall'AI o da un essere umano⁷. Come sostenuto⁸, i consumatori hanno il "diritto di sapere" se il contenuto che utilizzano è stato prodotto da una persona o da una macchina, in un possibile scenario in cui un consumatore riceva un testo come consiglio dal suo medico sulla sua salute. In quel caso, il consumatore vorrebbe sapere come è stato prodotto il parere e, nel caso in cui sia stato generato dall'AI, se è stato accuratamente controllato, dato il potenziale di errori negli output dei modelli GAI, e anche che l'eccessiva fiducia in tali modelli può creare seri problemi etici nell'assistenza sanitaria⁹.

Ciò che rende questo argomento di grande rilevanza è anche il fatto che, accanto ai rapidi sviluppi della tecnologia con l'emergere di nuovi sistemi e applicazioni, la GAI ha acquisito un ruolo centrale nel processo normativo che ha portato all'accordo all'interno del cosiddetto Trilogo (Commissione, Consiglio e Parlamento) per l'approvazione della legge sull'intelligenza artificiale (AI Act), che è stata approvata dal Parlamento il 13 marzo 2024. L'ascesa della ChatGPT ha costretto i politici dell'UE a riprendere i negoziati per concordare le modalità di inclusione della GAI nell'AI Act. Il modo in cui la GAI ha interrotto l'iter legislativo della legge sull'intelligenza artificiale è un esempio di quanto sia difficile per la regolamentazione essere "inossidabile" in un settore innovativo e dinamico come quello dell'intelligenza artificiale. Basta leggere il "Legislative Train Schedule" (uno strumento online prodotto dal Parlamento europeo per monitorare i processi legislativi) sulla legge sull'intelligenza artificiale per vedere che la Commissione europea ha presentato una proposta per una nuova legge sull'intelligenza artificiale (AI Act) nell'aprile 2021 e il Consiglio

ha adottato la sua posizione comune ("approccio generale") sulla legge sull'intelligenza artificiale il 6 dicembre 2022. Nessuna delle due proposte contemplava l'AI generativa. Ma il boom delle ChatGPT ha spinto il Parlamento ad assumere, a maggio 2023, una nuova posizione che richiede che le GAI siano regolate dall'AI Act¹⁰. Nella sua posizione, il Parlamento ha introdotto modifiche per garantire che l'AI Act disciplinasse l'AI generativa¹¹. Dopo un lungo dibattito, in data venerdì 8 dicembre 2023, i negoziatori del Parlamento e del Consiglio hanno raggiunto un accordo sulla legge sull'intelligenza artificiale. Un nuovo testo è stato pubblicato a gennaio 2024 e approvato dai rappresentanti degli Stati membri il 2 febbraio 2024 e dalle commissioni Mercato interno e Libertà civili del Parlamento europeo il 13 febbraio 2024. Il 13 marzo 2024 il testo è stato approvato in maniera definitiva dal Parlamento Europeo, per poi entrare ufficialmente in vigore, dopo la pubblicazione sulla Gazzetta Ufficiale Europea, il 2 agosto 2024. Oltre al lavoro legislativo sull'AI nell'UE, anche gli Stati Uniti si stanno muovendo rapidamente con vari mezzi tra cui l'ordine esecutivo Biden-Harris sull'AI. E anche il Regno Unito ha dato un impulso in tal senso mediante il recente vertice sulla sicurezza dell'AI nel Regno Unito. Ma la regolamentazione è solo uno dei tasselli di un più ampio quadro di governance sull'AI: gli altri sono gli investimenti pubblici, la politica della concorrenza e gli antitrust. A gennaio 2024, ad esempio, l'UE ha presentato un pacchetto sull'innovazione per sostenere le PMI e le start-up dell'AI, prevedendo la mobilitazione di investimenti pubblici affinché il messaggio fosse che l'UE non si limita a regolamentare (AI Act) ma che sostiene anche l'innovazione. Le recenti operazioni di mercato nel settore GAI (ad esempio, la partnership tra Mistral e Microsoft, si veda infra, sezione 3.2) hanno recentemente sollevato preoccupazioni sul potere di mercato e sulla necessità di garantire la concorrenza e mantenere il mercato GAI emergente aperto come un campo di gioco equo.

Questo report si basa su fonti secondarie, ragionamenti analitici e teorici, interviste selezionate e sapere di esperti. Una prima versione di questo report è stata discussa durante due workshop

con circa 20 esperti il 6 e 7 marzo 2024. Nel Capitolo 2 il report fornisce un'analisi sintetica di cosa sia e come funzioni la GAI. Nel capitolo 3 vengono presentate le dinamiche socio-economiche della GAI. Il capitolo 4 si concentra sui rischi, sul processo normativo e sul dibattito in corso. Il capitolo 5 presenta la prospettiva dell'UE. Nel Capitolo 6 vengono presentati quattro possibili scenari futuri per lo sviluppo dell'AI generativa, valutati successivamente nel Capitolo 7, e si presentano le implicazioni e le raccomandazioni di carattere politico.

COS'È E COME FUNZIONA L'AI GENERATIVA

Le grandi reti neurali che consentono il funzionamento della maggior parte delle GAI moderne sono chiamate modelli di base. Con qualche semplificazione, i modelli di base si distinguono dalle precedenti generazioni di modelli di intelligenza artificiale per il fatto che possono essere utilizzati per una serie di compiti precedentemente non specificati (mentre i modelli precedenti tendevano a essere "limitati" e potevano eseguire solo i compiti per i quali erano stati addestrati) e possono essere perfezionati per compiti e contesti diversi con un addestramento relativamente limitato. La dimensione di questi modelli è tipicamente misurata in numero di parametri (cioè i pesi delle connessioni tra i neuroni artificiali della rete) o di numero di passi computazionali (operazioni in virgola mobile o FLOP) necessari per addestrare il modello. Le reti neurali più grandi e all'avanguardia sono chiamate modelli di frontiera.

I trasformatori¹² sono un tipo di architettura di rete neurale che ha svolto un ruolo fondamentale nei recenti progressi dell'apprendimento automatico¹³ e che funge da spina dorsale della maggior parte dei modelli di frontiera. Il termine "trasformatore" riflette un approccio generalizzato alla trasformazione di una sequenza di input in una sequenza di output, come speech-to-text, text-to-speech, text-to-text e altre modalità di dati (sequenziabili) attraverso un meccanismo chiamato "attenzione". I trasformatori, insieme agli algoritmi di self-supervised learning che non richiedono l'impegno umano per l'etichettatura di grandi insiemi di dati, si sono dimostrati molto efficaci nell'elaborazione di informazioni sequenziali, rendendoli particolarmente adatti a compiti di elaborazione del linguaggio naturale come la generazione di testi, il riassunto e la risposta a domande. L'innovazione nell'uso

dei trasformatori per l'elaborazione del linguaggio naturale e altre applicazioni della GAI è parte integrante del successo dei modelli di base. Questi trasformatori consentono ai modelli di base di catturare e analizzare le relazioni contestuali tra molti elementi diversi in una sequenza, che si tratti di parole in una frase o di punti di dati in una serie temporale. Questa capacità è fondamentale per generare risultati coerenti e contestualmente rilevanti in varie applicazioni. Nel contesto dell'AI generativa, i trasformatori contribuiscono alla capacità dei modelli di base di apprendere e generare contenuti diversi in un'ampia gamma di compiti. Il Generative Pre-trained Transformer (GPT) è un esempio notevole di un modello di base che sfrutta i trasformatori, mostrando la loro importanza nel panorama della moderna intelligenza artificiale.

Finora, l'applicazione di GAI che ha ricevuto maggiore attenzione è ChatGPT, il chatbot di OpenAI che genera testo e risponde a domande e che utilizza i Large Language Models (LLM) GPT-3.5 e GPT-4. I LLM sono progettati per interpretare e generare testi simili a quelli umani sulla base dei modelli e delle informazioni apprese da vaste quantità di dati di addestramento. Questi modelli, compreso il GPT, sono pre-addestrati su enormi serie di dati contenenti testi diversi, che consentono di cogliere le sfumature del linguaggio, della grammatica e del contesto. L'espressione "LLM" sottolinea l'importanza di questi modelli linguistici nel campo della GAI, evidenziando la loro capacità di comprendere e generare testi coerenti e contestualmente rilevanti in diversi compiti. Questi modelli hanno dimostrato notevoli capacità in compiti di linguaggio naturale come il completamento di testi, la risposta a domande e la generazione di contenuti, rendendoli parte integrante dei progressi delle applicazioni di AI basate sul linguaggio.

Le soluzioni di GAI attualmente più diffuse si basano su due risorse essenziali: i dati e le risorse di calcolo. I dati e la potenza di calcolo rappresentano i costi. Le prestazioni delle GAI e i costi associati dipendono dal numero di parametri addestrabili¹⁴: più il numero di parametri aumenta più i modelli possono imparare, ma aumentano anche i costi in termini di dati e di potenza di calcolo richiesta. I costi in termini di dati e di potenza di calcolo sono sostenuti sia nella fase di addestramento sia in quella di inferenza. A questo proposito, va sottolineato che la comunità di ricerca sta sviluppando gli Small Language Models (SLM) che richiedono un minor numero di parametri addestrabili e mirano a raggiungere le stesse prestazioni dei LLM¹⁵ o in alternativa lo sviluppo di nuovi modelli attraverso la messa a punto di quelli esistenti utilizzando la tecnica Low-Rank Adaptation (LoRA)¹⁶; sia gli SLM che i LoRA richiedono meno dati e meno potenza di calcolo rispetto ai LLM. I dati utilizzati dai LLM devono essere vasti, diversificati, raccolti e aggiornati rapidamente (freschezza) e di qualità. Le prestazioni dei modelli aumentano in funzione del volume, della varietà e della qualità dei dati utilizzati per l'addestramento. A questo proposito, vale la pena sottolineare che i ricercatori hanno sviluppato un modello chiamato KOALA che, utilizzando dati di alta qualità ma con un volume inferiore (ha richiesto solo sei ore di addestramento), ha ottenuto prestazioni simili a quelle dei LLM che utilizzano volumi di dati molto più grandi¹⁷. Le risorse informatiche comprendono hardware di elaborazione, server, supercomputer e apparecchiature di rete. La potenza di calcolo può essere noleggiata da fornitori di cloud, il che in genere funziona bene per la messa a punto e per i modelli più piccoli, ma diventa troppo costoso per creare modelli di base da zero su larga scala. I LLM hanno bisogno di grandi quantità di potenza di calcolo per l'addestramento e il funzionamento. In particolare, hanno bisogno di unità di elaborazione grafica (GPU) per eseguire diversi calcoli simultaneamente. Intel detiene il 68,7% del mercato delle CPU, mentre Nvidia detiene l'82% del mercato delle GPU¹⁸. Si può quindi affermare che i mercati che forniscono risorse informatiche per i LLM sono piuttosto concentrati. D'altra parte, per la messa a punto e i modelli di piccole dimensioni, i fornitori di cloud affittano agli sviluppatori hardware di elaborazione, storage,

server e tecnologie di supercalcolo. Tra i grandi fornitori di questi servizi figurano Google (Google Cloud Platform), Amazon (Amazon Web Services) e Microsoft (Microsoft Azure). Vale la pena notare che alcuni di questi fornitori sono anche coinvolti nello sviluppo di LLM, sviluppando modelli propri o investendo in aziende che creano questi modelli come Mistral (ovvero Microsoft, si veda infra) e Anthropic e, quindi, hanno un vantaggio competitivo e un interesse commerciale nei modelli che funzionano su vasta scala. Il mercato dei cloud provider, tuttavia, è ancora competitivo e non concentrato¹⁹.

I LLM presentano due aspetti che suggeriscono un impatto potenzialmente rapido e diffuso sul lavoro. In primo luogo, capacità sorprendenti, poiché i LLM, originariamente progettati come modelli generali, manifestano conoscenze e abilità specialistiche che si evolvono con il miglioramento delle dimensioni e della qualità del modello²⁰. La GAI mette in mostra capacità nuove e impreviste con un'ampia applicabilità. Studi recenti evidenziano le loro prestazioni di alto livello in contesti professionali²¹. In secondo luogo, il miglioramento diretto delle prestazioni, perché i LLM hanno la capacità unica di migliorare direttamente le prestazioni degli utenti senza richiedere investimenti organizzativi o tecnologici sostanziali. Le prime ricerche sull'ultima generazione di LLM indicano un aumento tangibile delle prestazioni, in particolare nei compiti di scrittura²² e di programmazione²³, oltre che nel lavoro creativo e di ideazione.

I LLM, tuttavia, nel loro attuale stato di sviluppo non sono privi di difetti, caratterizzati da una relativa opacità, che comprende non solo il funzionamento interno dei modelli di AI, ma anche i loro punti deboli. Esistono numerose prove degli errori contenuti nei risultati dei modelli di GAI a causa delle limitazioni dei LLM sottostanti²⁴, generalmente definite "allucinazioni". Secondo Mons²⁵, tali allucinazioni sono in gran parte dovute a dati insufficienti e a «una mancanza di modelli concettuali convalidati per vincolare gli algoritmi e i risultati del LLM»²⁶. Mons sostiene che si dovrebbero distinguere tre categorie di input con cui i

modelli vengono alimentati: dati sperimentali, osservazioni del mondo reale e conoscenze consolidate, ma in realtà le “macchine” vengono alimentate come se tutti gli input fossero conoscenze consolidate. Conclude il suo ragionamento affermando che «Senza un modello concettuale che vincoli l’analisi, ci troviamo di fronte alla probabilità di molte allucinazioni, qui definite come schemi che hanno perfettamente senso per una macchina senza alcun vincolo basato su un modello concettuale, ma che non contengono alcuna conoscenza perseguibile per le persone»²⁷. Questo aspetto e gli altri svantaggi sono affrontati nella Sezione 4.1 del Capitolo 4.

DINAMICHE SOCIOECONOMICHE DELL'AI GENERATIVA

SVILUPPI DEL MERCATO

Lo sviluppo della GAI ha subito un'impennata dopo il lancio di ChatGPT da parte di OpenAI nel novembre 2022. L'impatto è stato profondo e ha portato a scoperte preziose e a progressi trasformativi nelle applicazioni GAI. Il ciclo di vita, la frequenza di utilizzo e la rapida integrazione nelle applicazioni software esistenti sono notevoli. Ad esempio, solo quattro mesi dopo il rilascio di ChatGPT, OpenAI ha introdotto GPT-4, basato su un nuovo LLM, che presenta capacità notevolmente migliorate. Poco dopo, Claude, la GAI di Anthropic, ha dimostrato un importante aumento della potenza di elaborazione, gestendo, a maggio 2023, 100.000 token di testo, equivalenti a circa 75.000 parole al minuto, rispetto alla sua introduzione a marzo 2023 con una capacità di circa 9.000 token.

Nel maggio 2023, Google ha annunciato una suite di nuove funzionalità basate sulla GAI, tra cui la Search Generative Experience e un LLM chiamato PaLM 2, destinato ad alimentare il suo chatbot Bard e altri prodotti Google. Nel dicembre 2023, Google ha annunciato Gemini, la serie di LLM più avanzata di Google. Il nuovo strumento è stato addestrato per essere multimodale, cioè in grado di elaborare diversi tipi di media come testo, immagini, video e audio. Google ha dichiarato che Gemini Ultra ha superato i modelli di intelligenza artificiale "all'avanguardia", tra cui il modello più potente di ChatGPT, GPT-4, in 30 dei 32 test di benchmark, tra cui test di ragionamento e di comprensione delle immagini. Altre aziende tecnologiche leader hanno rilasciato le proprie versioni di LLM, come Llama di Meta, dimostrando che il panorama dello sviluppo di modelli linguistici è entrato in un'era di intensa competizione.

È chiaro che le Big Tech sono ben posizionate per diventare potenzialmente dominanti, per varie ragioni, anche nel nuovo

dominio della GAI. Hanno accesso a molti dati, possiedono potenza di calcolo (e la affittano sotto forma di servizi cloud, generando potenzialmente dipendenze per gli operatori più piccoli), dispongono di molto denaro e hanno accesso ai migliori talenti.

L'ecosistema GAI è suddiviso in tre livelli principali: Fornitori di



Figura 1 Ecosistema di AI generativa, Fonte: Bornstein et al. (2023). Who Owns the Generative AI Platform? (<https://a16z.com/who-owns-the-generative-ai-platform/>)

applicazioni, modelli e infrastrutture²⁸ (si veda l'immagine sottostante)

Per analizzare gli sviluppi del mercato, è importante tenere a mente questo quadro. In primo luogo, i fornitori di modelli svolgono un ruolo fondamentale nel plasmare i progressi della GAI, semplicemente perché quest'ultima non esisterebbe senza il lavoro di ricerca e ingegneria svolto da aziende tecnologiche come Google, OpenAI, Amazon, Meta, X.AI, Stability e altri player, come i cinesi Tencent e Alibaba, ma anche dalla comunità accademica. Lo sviluppo iniziale dei modelli richiede un investimento significativo per via delle ampie risorse computazionali essenziali per l'addestramento e dello sforzo umano necessario per affinare tali modelli. Di conseguenza, questa fase iniziale di sviluppo è stata guidata prevalentemente da grandi colossi tecnologici, da start-up sostenute da ingenti investimenti e da selezionati collettivi di ricerca open-source (ad esempio BigScience). Tuttavia, gli sforzi attuali sono rivolti alla creazione di modelli più piccoli che possano dare risultati efficaci per compiti specifici e all'ottimizzazione dei processi di addestramento per maggiore efficienza, il che apre il campo anche alle start-up. Questo progresso, infatti, ha il potenziale per ampliare l'apertura del mercato, invitando nuovi partecipanti. In particolare, alcune start-up hanno già raggiunto il successo nello sviluppo indipendente dei propri modelli: ne sono un esempio Cohere, Anthropic, ElevenLabs e AI21 Labs, specializzate nella costruzione e nell'addestramento dei loro LLM. Esistono anche

player europei (ad es. Aleph Alpha, Poro, ecc.) che vengono presi in considerazione più avanti, nella Sezione 5.1.

Un'osservazione fondamentale per i fornitori di modelli è l'evidente connessione tra commercializzazione e hosting. La domanda di API di proprietà, esemplificata dalle offerte di OpenAI, è in rapida crescita. Allo stesso tempo, i servizi di hosting per i modelli open-source stanno emergendo come validi hub che facilitano la condivisione e l'integrazione dei modelli. Mentre i modelli costituiscono il nucleo cognitivo della GAI, un'intera catena del valore si sta evolvendo per supportare la formazione e l'utilizzo di questa tecnologia trasformativa. I fornitori di infrastrutture svolgono un ruolo fondamentale nel fornire la potenza di calcolo essenziale e necessaria per l'addestramento di questi modelli. Le piattaforme cloud, a loro volta, offrono l'accesso a questa infrastruttura hardware avanzata. Dominato dai principali player come AWS, GCP e Azure, con competitor emergenti, questo livello è alla base delle dinamiche finanziarie dell'AI generativa. Le stime suggeriscono che una parte significativa dei ricavi dell'AI generativa, tra il 10% e il 20%, passa attraverso le società di infrastrutture. Le start-up, sostenute da miliardi di capitale di rischio, contribuiscono a questo flusso finanziario, la maggior parte del quale viene speso per i fornitori di cloud. In particolare, Nvidia si distingue come player di punta, registrando un notevole fatturato per le GPU dei data center, derivante principalmente dai casi d'uso dell'intelligenza artificiale generativa. Tuttavia, sta emergendo una concorrenza attestabile, con aziende come Oracle e start-up come Coreweave e Lambda Labs che stanno guadagnando terreno grazie a soluzioni convenienti e a un'assistenza personalizzata. Va notato che la dipendenza da fornitori stranieri (in questo caso americani) per l'accesso all'infrastruttura cloud su cui vengono addestrati i modelli di Generative AI espone la Ue a rischi significativi, tra cui la perdita di controllo sui dati, la dipendenza tecnologica e i vincoli imposti dalle piattaforme proprietarie (vendor lock-in). Una soluzione potrebbe essere quella dell'adozione a livello europeo di un'infrastruttura cloud open source che, oltre a permettere alle imprese europee di avere un controllo più diretto sui propri sistemi digitali, potrebbe favorire un modello di innovazione collaborativa, aperto ad aziende di diverse

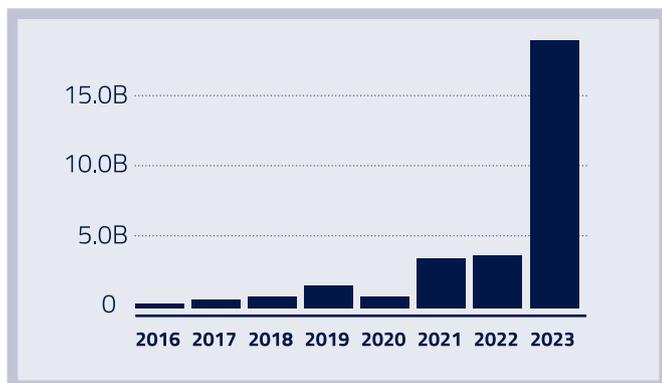


Figura 2 Investimenti globali nell'AI generativa, Fonte: Dealroom.co

dimensioni, evitando il rischio dell'eccessiva concentrazione di potere nelle mani di pochi grandi fornitori.

Inoltre, accanto a questo modello dominante basato sul cloud, a livello industriale cresce l'attenzione verso l'edge computing, in cui i dati vengono trattati in locale, prima di trasferirli sul cloud. I dati utilizzati dai modelli di intelligenza artificiale generativa provengono spesso da sensori distribuiti, e oggi i sensori devono non solo raccogliere informazioni, ma anche elaborare localmente parte del processo decisionale o generativo. Questa evoluzione avvicina l'AI alle fonti dei dati, migliorando l'efficienza operativa.

Per gestire questi cambiamenti, aziende come STMicroelectronics stanno esplorando l'uso di modelli di linguaggio più piccoli e specifici (Short Language Models, SLM), che potrebbero potenziare l'edge computing. Tuttavia, questo approccio richiede un significativo impegno in ricerca e sviluppo per comprendere come tali modelli possano gestire non solo la previsione, ma anche la generazione di contenuti in tempo reale, che è una delle caratteristiche distintive dell'AI generativa. La sfida è ulteriormente complicata dal fatto che l'hardware deve supportare un software in continua evoluzione, il che rende necessaria la progettazione di un'infrastruttura hardware flessibile, capace di adattarsi ai rapidi cambiamenti dei sistemi software.

Infine, le applicazioni integrano i modelli di intelligenza artificiale generativa a prodotti incentrati sull'utente. Che si tratti di orchestrare pipeline end-to-end o di sfruttare API esterne, queste applicazioni coprono ambiti diversi come la generazione di immagini, il copywriting e la creazione di codice. In particolare, numerose aziende si stanno affacciando sul mercato, presentando applicazioni costruite su modelli di base personalizzati per l'esecuzione di compiti specifici.

Insieme alla rapida evoluzione delle capacità di GAI, i finanziamenti per questa tecnologia trasformativa sono diventati una forza significativa nel panorama degli investimenti. Nei primi dieci mesi del 2023, la GAI ha raccolto 17,8 miliardi di dollari, un investimento

consistente e in rapida crescita²⁹. Gli investimenti in capitale di rischio e altri investimenti privati esterni in GAI hanno registrato un notevole tasso di crescita medio composto del 74% annuo dal 2017 al 2022, superando l'aumento del 29% annuo degli investimenti complessivi in intelligenza artificiale nello stesso periodo.

Guardando ai segmenti chiave della GAI, gli sviluppatori di modelli sono emersi come i beneficiari predominanti in termini di investimenti, assicurandosi oltre il 70% dei finanziamenti totali tra il 2019 e il 2023. Le applicazioni hanno ricevuto circa il 24% degli investimenti totali, evidenziando il ruolo centrale del software e delle implementazioni pratiche. Nel frattempo, le infrastrutture, essenziali per sostenere la spina dorsale di questi modelli avanzati, rappresentano il 5% dell'investimento complessivo, sottolineando il riconoscimento del ruolo critico svolto dalle basi tecnologiche nel far progredire le capacità della GAI. Questa distribuzione sottolinea le diverse aree di interesse all'interno dell'ecosistema, ognuna delle quali contribuisce alla continua crescita e allo sviluppo di questa tecnologia trasformativa.

L'aumento sostanziale degli investimenti nella GAI per tutto il 2023 non è solo un segno di entusiasmo, ma riflette una convinzione radicata tra gli investitori sulle implicazioni tangibili di questa tecnologia.

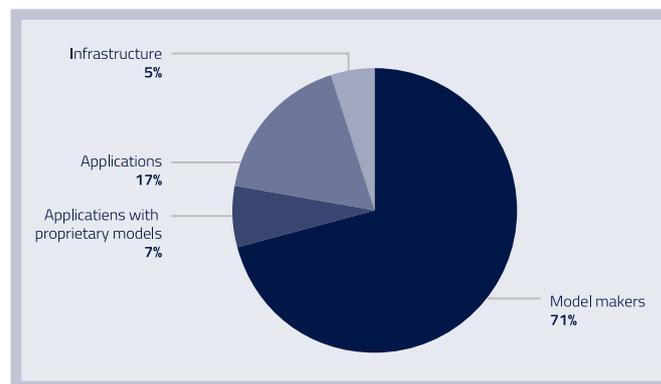


Figura 3 Finanziamenti in capitale di rischio per l'AI generativa (2019-2023), divisi per segmento, Fonte: Dealroom.co

L'aumento del sostegno finanziario è indicativo dell'aspettativa condivisa che l'AI generativa modificherà radicalmente il nostro approccio al lavoro e migliorerà la produttività complessiva. Gli investitori stanno stanziando fondi significativi nell'attesa di assistere a risultati concreti e significativi da parte della GAI. Guardando al 2024, l'Economist prevede un notevole spostamento dell'attenzione verso le aziende al di fuori del settore tecnologico, che abbracceranno la GAI³⁰. I loro obiettivi principali saranno la riduzione dei costi e l'aumento della produttività. A sostegno di questa aspettativa ci sono diversi fattori convincenti, che rappresentano un punto di svolta cruciale nell'adozione diffusa dell'AI nel panorama aziendale. Mentre nel 2023 le principali aziende hanno esplorato la GAI, sperimentando applicazioni che vanno dalla stesura di documenti all'analisi di trascrizioni, nel 2024 le aziende si stanno preparando a una maggiore diffusione. Un'indagine di KPMG ha rivelato che oltre l'80% delle aziende statunitensi prevede di aumentare gli investimenti in GAI di oltre il 50% entro la metà del 2024.³¹

Esplorando il panorama dei ricavi dell'ecosistema dell'AI, possiamo osservare un modello diverso. In seguito all'introduzione di ChatGPT nel 2022, OpenAI ha registrato una notevole impennata in un anno, raggiungendo un traguardo di fatturato senza precedenti, passando da zero a 1 miliardo di dollari, secondo solo all'impressionante performance del produttore di microchip statunitense NVIDIA³². Le vendite di GPU per data center di NVIDIA sono passate da 3,6 miliardi di dollari nel quarto trimestre del 2022 a 16 miliardi di dollari nel quarto trimestre del 2023, consolidando la sua posizione di leader nello spazio dell'AI generativa. In termini di fatturato, il segmento di mercato delle infrastrutture dati è in testa. Il mercato ha registrato una crescita straordinaria, raggiungendo i 49 miliardi di dollari nel 2023, con un sostanziale aumento del 182% rispetto all'anno precedente. NVIDIA è emersa come player dominante, con un'impressionante quota di mercato pari al 92%. Nonostante le sfide, come l'aumento dei prezzi e i vincoli di fornitura, la domanda di GPU per data center rimane solida. Nel 2023 il fatturato trimestrale di NVIDIA ha registrato un aumento sbalorditivo del 272%, passando da 4,3 miliardi di dollari nel primo trimestre a 16 miliardi di dollari nel quarto trimestre. Il secondo

segmento di mercato, in termini di fatturato, è composto da aziende di modelli di base. Questo mercato ha subito una trasformazione con il rilascio pubblico di ChatGPT alla fine del 2022, raggiungendo un fatturato di 3,0 miliardi di dollari nel 2023. Le proiezioni indicano un futuro promettente, con le imprese che investono massicciamente in implementazioni e applicazioni di AI generativa. OpenAI domina questo mercato con una quota del 39%, seguita da vicino da Microsoft con il 30%. Azure AI di Microsoft, che sfrutta i LLM di OpenAI, si distingue per la maggiore sicurezza dei dati e le applicazioni AI personalizzabili. AWS, con una quota di mercato dell'8%, ha introdotto il servizio Bedrock nel settembre 2023, che fornisce l'accesso ai modelli di varie aziende di AI, tra cui Anthropic, AI21 labs e Cohere, ciascuna con una quota di mercato del 2%. Bedrock combina questi modelli con set di strumenti per sviluppatori, facilitando la costruzione e la scalabilità di applicazioni di AI generativa. Il mercato dei modelli di base e delle piattaforme continua a evolversi e si prevede una crescita significativa nei prossimi anni.

PROBLEMI DI CONCORRENZA

Sebbene la maggior parte degli sforzi dei governi riguardi la regolamentazione generale dell'AI con qualche riferimento specifico alla GAI, come nel caso dell'iniziativa UE dell'AI Act, il quadro di governance dell'AI e della GAI comprende anche una parte di politica della concorrenza e l'antitrust. Sebbene la GAI sia emersa come nuovo segmento di mercato solo negli ultimi due anni, il rapido ritmo di sviluppo ha già sollevato preoccupazioni circa la garanzia di concorrenza e di condizioni di parità. Le Big Tech, infatti, possiedono alcuni vantaggi che potrebbero portarle a dominare il nuovo mercato della GAI. Molto recentemente, alla fine di febbraio 2024, l'accordo di Microsoft con la start-up tecnologica francese Mistral AI, ad esempio, ha suscitato proteste nell'Unione Europea, con i legislatori che hanno chiesto un'indagine su quella che considerano una concentrazione di potere da parte del gigante tecnologico³³. Le autorità antitrust stanno già esaminando la partnership tra Microsoft e il produttore di ChatGPT OpenAI, e la Commissione europea che ha avvertito in precedenza che il rapporto tra le due società potrebbe violare le regole della concorrenza dell'UE. Quando alla fine dell'anno

scorso ha dovuto affrontare le pressioni normative per il suo investimento multimiliardario in OpenAI, Microsoft ha cercato di chiarire che in realtà non possedeva una quota della società e quindi non poteva controllarla. Tuttavia, Microsoft ha ottenuto l'accesso ai modelli all'avanguardia di OpenAI e il diritto di condividere i profitti di OpenAI in cambio dell'investimento, che è sotto esame antitrust sia nell'Unione Europea sia negli Stati Uniti.

Nel Regno Unito, la Competition and Markets Authority (CMA) ha avviato una revisione dei modelli di base (FM) di AI per capire come funziona il mercato³⁴. In particolare, la CMA esamina le barriere che i FM riscontrano per accedere al mercato e il loro impatto sulla concorrenza in altri mercati. Ma è soprattutto negli Stati Uniti che le agenzie di protezione dei consumatori e la Federal Trade Commission (FTC), in particolare, hanno espresso serie preoccupazioni sul rischio di concorrenza³⁵. La presidente della FTC, Lina Khan, ha recentemente messo in guardia dai notiziari nazionali sul fatto che «L'AI potrebbe essere usata per aumentare le frodi e le truffe» e ha aggiunto che la FTC sta vigilando affinché le grandi aziende non usino l'AI per «schiacciare la concorrenza»³⁶. La FTC ha prodotto diverse analisi e di seguito riportiamo i contenuti principali di un blog intitolato Generative AI Raises Competition Concerns³⁷.

Secondo il blog della FTC, sono tre le questioni principali che possono ridurre la concorrenza, cui si aggiunge la possibilità che i player adottino metodi di concorrenza sleale. In primo luogo, il controllo sui dati può creare barriere all'ingresso a vantaggio dei grandi player che possiedono anche piattaforme digitali e hanno già accumulato grandi quantità di dati. Secondo la FTC, quindi, il volume e la qualità dei dati necessari per pre-addestrare un modello di GAI da zero possono influire sulla capacità dei nuovi player di entrare nel mercato. In secondo luogo, un requisito fondamentale per costruire il modello di GAI è il talento. Siccome questo tipo di talento è scarso, le aziende grandi e floride sono avvantaggiate e possono finire per escludere i nuovi arrivati dall'accesso al talento necessario. In terzo luogo, l'accesso alla potenza di calcolo richiesta potrebbe anche diventare una barriera all'ingresso a favore delle grandi aziende ricche di risorse

finanziarie e/o che possiedono una propria potenza di calcolo e un'offerta di servizi cloud. Se da un lato la FTC riconosce che i piccoli nuovi player possono semplicemente affittare la potenza di calcolo dei servizi cloud, dall'altro sottolinea che tali servizi sono costosi e potrebbero diventarlo ancora di più in futuro, oltre a creare dei lock-in, il che ci porta a quello che consideriamo l'elemento più rilevante dell'analisi della FTC, ossia il rischio di possibili metodi di concorrenza sleale. I grandi player possono impegnarsi in operazioni di bundling e/o di vendita abbinata. Il bundling si verifica quando le aziende offrono più prodotti in un unico pacchetto. La FTC prevede quindi una serie di rischi che vale la pena riportare integralmente qui di seguito:

Gli operatori storici che controllano fattori di produzione chiave o mercati adiacenti, tra cui quello del cloud computing, potrebbero utilizzare metodi di concorrenza sleale per consolidare il loro attuale potere o utilizzarlo per ottenere il controllo su un nuovo mercato dell'AI generativa.

[...]

Gli operatori storici potrebbero essere in grado di collegare le nuove applicazioni di AI generativa con i prodotti principali esistenti per ridurre il valore delle offerte di AI generativa autonome dei concorrenti con potenziali distorsioni della concorrenza. Gli operatori storici che offrono una gamma di prodotti e servizi nell'ambito di un ecosistema possono anche adottare comportamenti esclusivi o discriminatori, indirizzando gli utenti verso i propri prodotti di AI generativa anziché verso quelli dei concorrenti. Inoltre, gli operatori storici che offrono sia servizi di calcolo che prodotti di AI generativa, ad esempio attraverso partnership esclusive nel cloud, potrebbero sfruttare il loro potere nel settore dei servizi di calcolo per soffocare la concorrenza nell'AI generativa, riservando a loro stessi e ai loro partner un trattamento discriminatorio rispetto ai nuovi operatori. Uno scenario correlato è quello in cui un operatore storico offre sia prodotti propri che sfruttano l'AI generativa, sia API che consentono ad altre aziende di sfruttare le proprie capacità di AI generativa. In tali circostanze, c'è il rischio che le imprese già presenti offrano le loro API a condizioni che servono a proteggere

la loro posizione dominante. Le aziende storiche potrebbero anche utilizzare le attività di fusione e acquisizione nello spazio dell'AI generativa per consolidare il potere di mercato nelle mani di pochi player. Le grandi aziende già attive nell'AI generativa, o che già controllano un input critico, potrebbero cercare di acquisire numerose applicazioni fondamentali e impedire l'accesso ai prodotti cardine da parte dei concorrenti. I leader di mercato possono anche cercare di acquistare applicazioni complementari e raggrupparle insieme. Inoltre, gli operatori storici possono essere tentati di acquistare semplicemente i rivali nascenti invece di cercare di superarli offrendo prodotti o servizi migliori.³⁹

Un quadro in qualche modo meno pessimistico si trova in un recente report di Bruegel³⁹, secondo il quale i mercati dei modelli di GAI sono attualmente ancora competitivi, con molti fornitori e un discreto grado di apertura. Le barriere all'ingresso non sono ancora insormontabili, sebbene anche questa relazione sottolinei alla fine lo stesso tipo di rischio di metodi di concorrenza sleale illustrato dalla FTC. Secondo questo report, i mercati dei LLM sono competitivi e dinamici, con costanti lanci di nuovi modelli di diverse dimensioni e con vari gradi di apertura da parte di una grande varietà di sviluppatori. Due fattori possono ridurre la necessità di risorse finanziarie per l'addestramento e l'esecuzione dei modelli: in primo luogo, lo sviluppo degli SLM (Small Language Models) e delle tecniche di messa a punto; in secondo luogo, la maggiore disponibilità di una comunità crescente di sviluppatori e di ricerca open-access che può ridurre le risorse umane necessarie. A livello di accesso ai dati, alcuni operatori (ad es. Google) traggono vantaggio dall'aver dati propri per addestrare ed eseguire i modelli. Tuttavia, gli sviluppatori possono utilizzare il numero crescente di dataset open-source disponibili sul web⁴⁰. Inoltre, in Europa gli sviluppatori che desiderano accedere ai dati possono invocare il Digital Market Act [DMA, Regolamento (UE) 2022/1925, articolo 6(11)], che impone agli operatori, identificati come "gatekeeper", l'obbligo di condividere i dati di ricerca (ranking, query, click e dati di visualizzazione). Quindi, secondo il rapporto di Bruegel, le barriere all'ingresso a livello di accesso ai dati non sono

attualmente proibitive. A livello di potenza di calcolo, si riconosce che i player che offrono potenza di calcolo/servizi cloud e sviluppano il proprio modello sono avvantaggiati. I grandi fornitori di cloud potrebbero godere di un vantaggio competitivo grazie alle economie di scala, all'accesso a grandi risorse finanziarie e alla fornitura di modelli propri esclusivamente sui loro servizi cloud. Il mercato dei servizi cloud, tuttavia, è ancora competitivo e offre ai piccoli sviluppatori la possibilità di affittare la potenza di calcolo necessaria in modo scalabile per la messa a punto e il lavoro con modelli più piccoli, senza la necessità di grandi investimenti finanziari iniziali. Anche in questo settore, lo sviluppo di SLM e di tecniche di messa a punto ridurrà la necessità di potenza di calcolo. Pertanto, il rapporto di Bruegel conclude affermando che anche in questo caso le barriere all'ingresso sono ridotte.

Infine, c'è la questione dell'open-source che può avere conseguenze ambivalenti e incerte. Molto rapidamente sono emerse le capacità di generazione di immagini open-source che hanno eclissato quelle dei modelli di base propri che le hanno ispirate. In linea di principio, l'open-source potrebbe aprire il campo da gioco, una volta che i modelli di base diventano pubblici. Tuttavia, c'è anche il rischio che l'open-source venga utilizzato in modo improprio. Immaginiamo uno strumento open-source per la generazione di immagini AI rilasciato con restrizioni incorporate sui tipi di immagini che possono essere generate. È possibile che utenti malintenzionati rimuovano queste protezioni e utilizzino i modelli per creare immagini private non consensuali. Inoltre, in passato abbiamo assistito alla pratica "apertura prima, chiusura poi". Anticipando qui la discussione sulla legge europea sull'AI, questa ambivalenza dell'open-source spiega perché il Parlamento Europeo non è stato pienamente soddisfatto dell'esenzione concessa ai modelli open-source contenuta nell'ultimo compromesso sulla legge sull'AI concordato nel dicembre 2023 (cfr. infra, sezione 4.3).

IMPATTI ECONOMICI

Le generazioni precedenti di tecnologie di automazione eccellevano nell'automatizzare le attività di gestione dei dati legate alla raccolta

e all'elaborazione delle informazioni. La GAI, con le sue capacità di linguaggio naturale, aumenta il potenziale di automazione di queste attività cognitive. L'influenza più significativa della GAI è prevista nel lavoro di conoscenza, in particolare nelle attività che richiedono processi decisionali e di collaborazione, che storicamente avevano un potenziale di automazione inferiore. Il forte aumento del potenziale di automazione è in gran parte attribuito alla capacità della GAI di comprendere e utilizzare il linguaggio naturale in varie attività e compiti. Questo aumento del potenziale di automazione implica che numerose attività lavorative che coinvolgono la comunicazione, la supervisione, la documentazione e l'interazione generale con le persone hanno il potenziale per essere automatizzate dalla GAI.

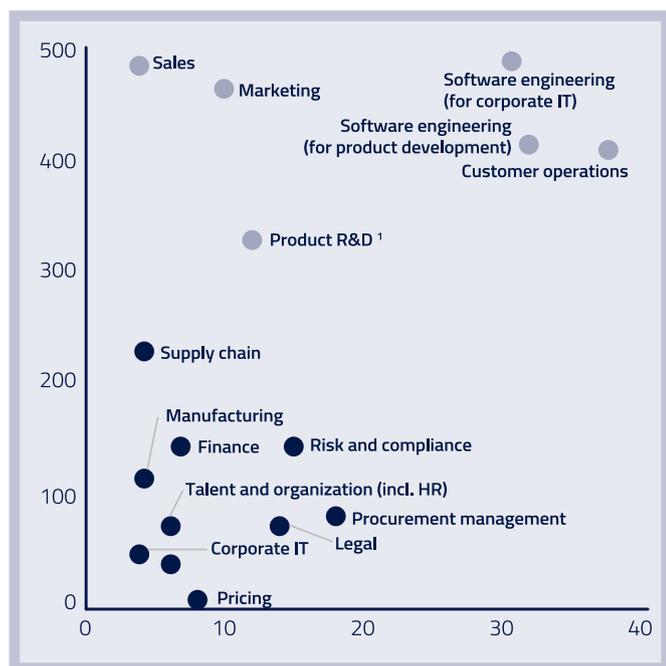


Figura 4 Impatto dell'AI generativa per funzione aziendale, Fonte: Chui et al (2023) si veda la nota a piè di pagina 41.

McKinsey ha condotto un'analisi approfondita dell'impatto potenziale della GAI sull'economia globale, esaminando le attività lavorative essenziali in circa 850 occupazioni⁴¹. L'analisi ha comportato la modellazione di scenari per stimare quando la GAI potrebbe eseguire con competenza oltre 2.100 attività lavorative specifiche, che vanno dalla "comunicazione con altri su piani o attività operative" ad altri compiti declinati all'interno di queste professioni nell'economia globale. Attraverso questo approccio, McKinsey ha misurato la potenziale influenza delle attuali capacità GAI sulla produttività del lavoro nell'intero spettro di attività svolte dalla forza lavoro globale. Il risultato rivela notevoli benefici economici, stimando un aumento annuale della produttività valutato tra i 6,1 e i 7,9 bilioni di dollari. Si tratta di un notevole aumento della produttività, compreso tra il 35% e il 70%, che sottolinea il potenziale di trasformazione della GAI su scala globale. Secondo l'analisi economica di McKinsey, delle 16 funzioni aziendali che saranno influenzate dall'AI, quattro (operazioni con i clienti, marketing e vendite, ingegneria del software e ricerca e sviluppo) potrebbero rappresentare circa il 75% del valore totale annuo dei casi d'uso dell'AI generativa. Pertanto, l'utilizzo della GAI in queste funzioni potrebbe determinare la maggior parte dell'impatto economico di questa tecnologia.

Una recente pubblicazione che analizza il potenziale impatto sul mercato del lavoro dei modelli LLM⁴² rivela che circa l'80% della forza lavoro statunitense potrebbe subire un impatto su almeno il 10% delle proprie mansioni lavorative a seguito dell'introduzione dei LLM, mentre circa il 19% dei lavoratori potrebbe subire un impatto su almeno il 50% delle proprie mansioni. Questo studio ha analizzato più di 1.000 occupazioni e 19.000 mansioni e gli autori osservano che la maggior parte delle occupazioni presenta un certo grado di esposizione ai LLM: le occupazioni a più alto salario in genere presentano un maggior numero di mansioni ad alta esposizione. In termini di competenze, in linea con la figura precedente tratta dal summenzionato report di McKinsey, i risultati dello studio indicano che le occupazioni che richiedono capacità di pensiero critico e quelle legate alla scienza hanno meno probabilità di essere influenzate dagli attuali LLM. Al contrario, le professioni che richiedono capacità

di programmazione e di scrittura sono più suscettibili a essere influenzate dai LLM. I recenti dati sperimentali sull'impatto della GAI sulla produttività sono rivelatori dell'impatto futuro di questa tecnologia. Come già accennato, i software engineer sembrano essere i lavoratori più colpiti dalla diffusione dell'AI generativa. I software engineer possono utilizzare l'AI generativa nella programmazione a coppie e per eseguire la codifica aumentata e addestrare i LLM a sviluppare applicazioni che generano codice quando viene fornito un prompt in linguaggio naturale che descrive ciò che il codice dovrebbe fare. Uno studio sperimentale ha utilizzato una prova controllata di GitHub Copilot, un programmatore di coppia di AI che suggerisce codice e intere funzioni in tempo reale in base al contesto. I risultati dello studio mostrano che il gruppo che ha utilizzato l'AI generativa ha completato il compito il 55,8% più velocemente rispetto al gruppo che non vi ha avuto accesso. Un altro interessante studio sperimentale analizza l'impatto della GAI sulla consulenza manageriale, un settore tradizionale ad alta intensità di conoscenza.

Utilizzando i dati di 758 consulenti del Boston Consulting Group, una società di consulenza manageriale globale, l'esperimento analizza le implicazioni dell'AI sulle prestazioni in compiti realistici, complessi e ad alto tasso di conoscenza⁴⁴. Lo studio dimostra che per ogni compito di consulenza realistico che rientra nello spettro delle capacità dell'AI, i consulenti che hanno utilizzato l'AI sono stati significativamente più produttivi (hanno completato in media il 12,2% in più di compiti, il 25,1% più rapidamente) e hanno prodotto risultati di qualità significativamente superiore (oltre il 40% di qualità in più rispetto al gruppo di controllo). Tuttavia, per un compito selezionato al di fuori dello spettro delle capacità dell'AI, i consulenti che utilizzano l'AI hanno avuto il 19% in meno di probabilità di produrre soluzioni corrette rispetto a quelli senza AI. Inoltre, come mostrato nella figura seguente, lo studio rivela che i lavoratori poco produttivi e alle prime armi tendono a beneficiare maggiormente dell'AI generativa rispetto ai lavoratori esperti e altamente qualificati. Questo risultato è confermato da un altro studio sperimentale, incentrato su un altro settore, che ha analizzato l'impatto di un assistente conversazionale

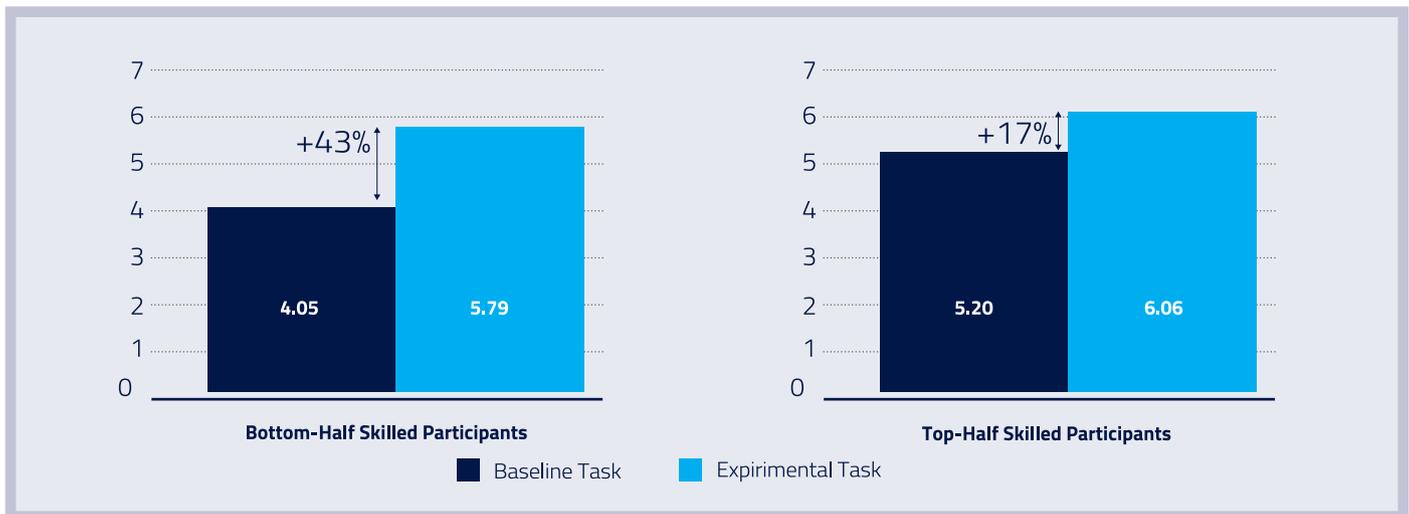


Figura 5 Impatto dell'AI generativa sui lavoratori della conoscenza poco e molto qualificati, Fonte: Dell'Acqua et al., (2023) si veda la nota a piè di pagina n. 45.

basato sull'AI generativa utilizzando i dati di 5.179 agenti di assistenza clienti⁴⁵. Sebbene lo studio dimostri che l'assistenza dell'AI aumenta complessivamente la produttività dei lavoratori, con un aumento del 14% del numero di chat che un agente risolve con successo all'ora, gli impatti sulla produttività dell'assistenza dell'AI sono molto disomogenei. I lavoratori meno qualificati e con meno esperienza migliorano significativamente in tutti i parametri di produttività. Al contrario, lo studio rileva un impatto minimo sulla produttività dei lavoratori più esperti o più qualificati.

La GAI potrebbe avere un impatto su un ampio numero di funzioni in diversi settori. La sua capacità di generare autonomamente contenuti, risolvere problemi complessi e imitare la creatività umana ha il potenziale per trasformare qualsiasi settore. Tuttavia, alcuni settori possono subire cambiamenti più profondi, poiché la versatilità della tecnologia promette di sbloccare nuove dimensioni di efficienza, creatività e risoluzione dei problemi. Come dimostrato da un recente studio, l'impatto complessivo dei LLM sul valore delle imprese è maggiore nei settori con occupazioni maggiormente esposte alla

GAI⁴⁶. Sembra tuttavia che l'impatto sia diffuso e ancora difficile da prevedere. Ad esempio, l'ultima indagine annuale McKinsey Global Survey sullo stato dell'AI mostra che la maggior parte delle persone in quasi tutti i settori del mondo ha provato uno strumento di AI almeno una volta nel 2023⁴⁷. In questi primi giorni, le aspettative per l'impatto dell'AI generativa sono elevate: tre quarti di tutti gli intervistati prevedono che l'AI generativa causerà cambiamenti significativi o dirompenti nella concorrenza del proprio settore nei prossimi tre anni. La figura precedente mostra la quota di intervistati che hanno utilizzato la GAI nel 2023 per settore. In generale, gli operatori del settore tecnologico, dei media e delle telecomunicazioni sono stati molto più propensi a utilizzare regolarmente gli strumenti di GAI. Ciò forse non sorprende, dato che questi settori sono i più propensi a coinvolgere l'uso della GAI nelle pratiche commerciali.

Come già evidenziato in precedenza, le perturbazioni provocate dalla GAI non sono uniformi in tutti i settori. I settori con attività ad alto tasso di conoscenza probabilmente assisteranno a un maggior numero di trasformazioni, con potenziali benefici sostanziali. Al contrario, le

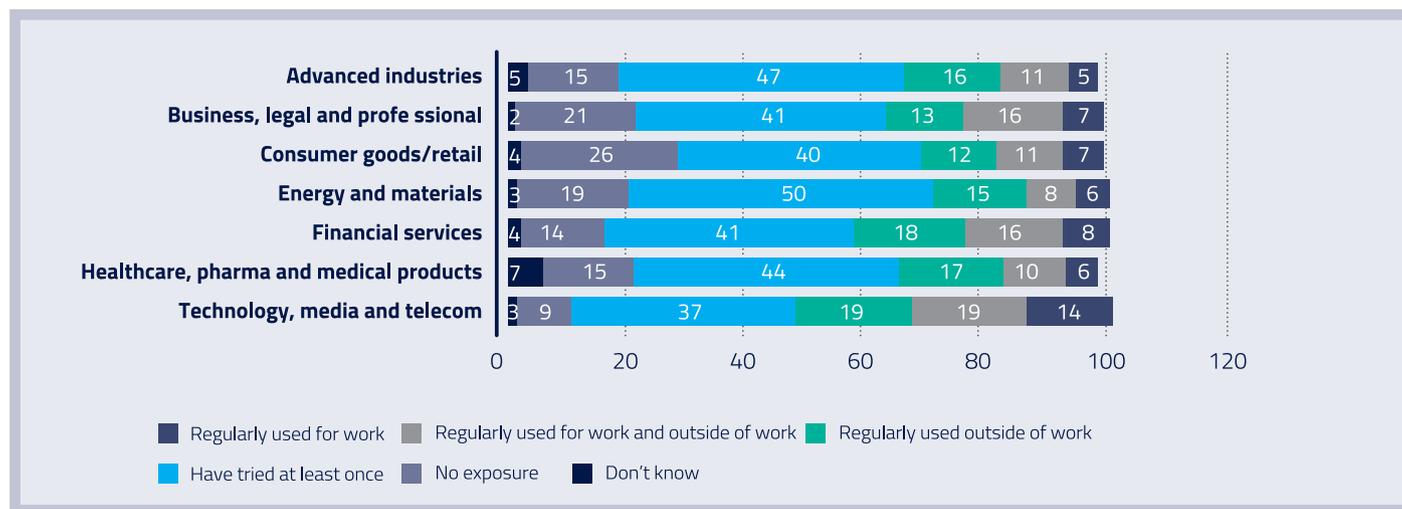


Figura 6 Quota di intervistati che utilizzano l'AI generativa nel 2023 per settore, Fonte: McKinsey & Company, si veda la nota 47

industrie radicate nel settore manifatturiero, come l'aerospaziale, l'automobilistica e l'elettronica avanzata, potrebbero subire un cambiamento relativamente minore. Questo allontanamento dallo schema storico delle ondate tecnologiche che hanno avuto un impatto prevalente sul settore manifatturiero deriva dalla capacità dell'AI generativa di svolgere attività basate sul linguaggio piuttosto che sul lavoro fisico. Di seguito ci concentreremo, a scopo puramente esemplificativo, su due settori specifici, per i quali discuteremo dei benefici e delle sfide presentate dalla GAI.

Nella parte restante di questa sezione approfondiremo due settori specifici, analizzando le sfumature con cui questa tecnologia può avere un impatto sui settori secondo diversi gradi di intensità. Iniziamo dal settore sanitario, per poi passare a quello energetico.

Assistenza sanitaria

L'avvento della GAI rappresenta una potenziale svolta per le

operazioni sanitarie, in particolare per quanto riguarda la gestione dell'abbondanza di dati non strutturati presenti in note cliniche, immagini diagnostiche, cartelle cliniche e registrazioni. Questi insiemi di dati non strutturati, utilizzati indipendentemente o insieme ad altri strutturati, come le richieste di risarcimento, rappresentano una frontiera promettente per le applicazioni trasformative in ambito sanitario. Oltre alla documentazione clinica, stanno emergendo diversi casi d'uso della GAI in ambito sanitario che generano sia entusiasmo sia preoccupazioni tra gli operatori sanitari. Nel settore sanitario, la salvaguardia delle informazioni sui pazienti è fondamentale dato che si tratta di informazioni altamente sensibili. A causa delle occasionali imprecisioni nelle risposte formulate dalla GAI, una stretta collaborazione e un attento monitoraggio da parte degli operatori sanitari sono essenziali per ridurre i rischi e garantire risultati affidabili. Mentre la tecnologia AI viene impiegata da anni nel settore sanitario per la previsione di eventi avversi e l'ottimizzazione della

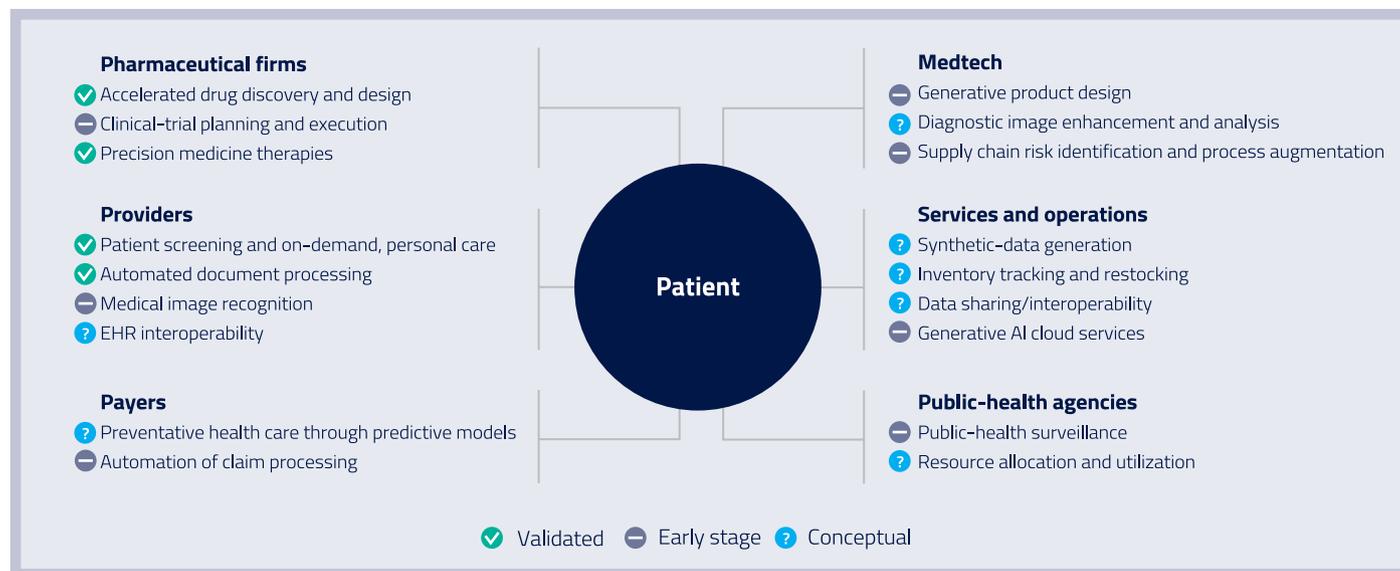


Figura 7 Potenziale caso d'uso in tutti i segmenti dell'assistenza sanitaria, Fonte: Boston Consulting Group

programmazione delle sale operatorie, la GAI può trasformare le nuove operazioni, sbloccando un importante valore potenziale del settore⁴⁸. Come mostra la figura seguente, estratta da un'analisi condotta dal Boston Consulting Group, la GAI genera numerosi casi d'uso e nuovi potenziali casi d'uso emergeranno nei prossimi anni.

In questa sezione, ci concentreremo sulle potenziali applicazioni in due aree specifiche: le operazioni cliniche e il processo decisionale clinico. In seguito tratteremo anche di alcuni dei principali rischi associati all'introduzione di questa tecnologia nel settore sanitario.

1. Operazioni cliniche. Le operazioni cliniche rappresentano un ambito promettente per le potenziali efficienze che la GAI potrebbe introdurre. Attualmente, gli operatori ospedalieri e il personale amministrativo sono alle prese con le numerose pratiche burocratiche associate all'assistenza dei pazienti, tra cui numerosi moduli, note post-visita e compiti amministrativi che contribuiscono a un notevole dispendio di tempo e a un potenziale esaurimento nervoso. Questo onere amministrativo è sentito anche dai gruppi di medici. La GAI, sotto la supervisione di un medico, promette di generare relazioni di dimissione o istruzioni nella lingua madre del paziente per una migliore comprensione. Inoltre, potrebbe sintetizzare le note di coordinamento dell'assistenza, le note di cambio turno e le liste di controllo in tempo reale, nonché i riepiloghi di laboratorio dei turni dei medici e gli ordini clinici. La capacità della tecnologia di generare e sintetizzare il linguaggio rappresenta anche un'opportunità per migliorare le cartelle cliniche elettroniche, sostituendo i requisiti di inserimento manuale e riducendo al minimo il rischio di errore umano. Gli ospedali e i gruppi di medici stanno esplorando attivamente le applicazioni di AI generativa, che vanno dalla pre-compilazione dei riepiloghi delle visite nelle cartelle cliniche elettroniche al perfezionamento della documentazione e alla fornitura di ricerche per il supporto alle decisioni. Alcuni sistemi sanitari hanno già incorporato questa tecnologia in programmi pilota. Il riquadro sottostante mostra un esempio di applicazione dell'AI generativa nelle operazioni cliniche. Riquadro 1 Esempio di applicazione pratica dell'AI generativa in ambito sanitario

La GAI è stato uno degli argomenti principali discussi alla Conferenza ed Expo annuali della HIMSS (Healthcare Information and Management Systems Society), uno dei più grandi e influenti eventi di informatica sanitaria al mondo. Una sessione si è concentrata su come la GAI abbia il potenziale per cambiare il modo in cui viene praticata la medicina. Con casi d'uso che vanno dall'inserimento ed elaborazione dei dati al supporto decisionale clinico e al coinvolgimento dei pazienti, la GAI offre molte opportunità per creare un sistema sanitario più efficiente, efficace e sostenibile, soprattutto se abbinata all'automazione. A dimostrazione delle sue capacità, la tecnologia ha mostrato come un operatore sanitario possa sfruttare senza problemi le nuove piattaforme per trasformare l'interazione con il paziente in note cliniche strutturate in pochi secondi. Utilizzando l'app mobile di AI platform, un medico può registrare la visita di un paziente. In tempo reale, la piattaforma incorpora le informazioni del paziente, identificando in modo intelligente tutti i dettagli mancanti e invitando il medico a completarli. In questo modo la dettatura iniziale si trasforma in una nota meticolosamente strutturata con un flusso conversazionale naturale. Dopo la visita, il medico accede agli appunti generati dall'intelligenza artificiale su un computer, dove può modificare il contenuto senza sforzo attraverso comandi vocali o digitando. Le note finalizzate vengono poi rapidamente inviate alla cartella clinica elettronica del paziente. Questo processo quasi istantaneo rende apparentemente arcaica l'annotazione manuale e dispendiosa in termini di tempo, un tempo obbligatoria per ogni interazione con il paziente.

Riquadro 1 Esempio di applicazione pratica dell'AI generativa in ambito sanitario

2. Processo decisionale clinico. Anche l'impatto dei LLM sul processo decisionale clinico dovrebbe essere significativo. La tecnologia è in grado di fornire un'interfaccia sofisticata per organizzare, recuperare e sintetizzare fatti, note e ricerche mediche complesse, rappresentando un cambiamento paradigmatico nel modo in cui i medici esaminano e interpretano i dati sanitari

complessi. Sfruttando le capacità dell'AI generativa, i medici possono potenzialmente semplificare le loro interazioni con grandi quantità di informazioni sui farmaci. Attraverso semplici interrogazioni, i medici possono sfruttare la potenza dei LLM per accedere a dati completi e aggiornati sui farmaci, migliorando la loro capacità di prendere decisioni informate sulla cura dei pazienti. Inoltre, l'AI generativa apre la strada a una maggiore personalizzazione dell'assistenza dei pazienti. Dando ai medici la possibilità di disporre di LLM, è possibile generare approfondimenti su misura, consentendo una comprensione più specifica delle esigenze dei singoli pazienti e delle opzioni terapeutiche. Sebbene l'integrazione della GAI nel processo decisionale clinico rappresenti un allontanamento dai flussi di lavoro tradizionali, i potenziali benefici in termini di migliore accesso alle informazioni, supporto decisionale potenziato e assistenza personalizzata al paziente sottolineano l'impatto trasformativo di questa tecnologia nel plasmare il futuro dell'assistenza sanitaria.

3. Rischi. Se da un lato l'adozione della GAI nell'assistenza sanitaria porta con sé diverse opportunità, dall'altro vi sono potenziali rischi da mitigare, soprattutto con la continua evoluzione della tecnologia. È necessario discutere alcune considerazioni cruciali. In primo luogo, vi sono problemi di sicurezza dei dati, in quanto il livello di sicurezza fornito dagli strumenti Gen-AI open-source può essere inferiore ai severi requisiti richiesti dal settore sanitario. Il rischio di violazione dei dati rappresenta una minaccia significativa che potrebbe compromettere le informazioni sensibili dei pazienti. In secondo luogo, il ricorso a dataset che sovraindicizzano determinate popolazioni di pazienti rappresenta un rischio sostanziale nella generazione dei piani di cura dei pazienti. Se i modelli sono addestrati su dati distorti, anche i piani di cura generati potrebbero essere sbagliati, fornendo ai pazienti informazioni imprecise, non utili o addirittura dannose. Per ridurre questo rischio è indispensabile garantire dati di formazione diversificati e rappresentativi. In terzo luogo, data la possibilità che l'AI produca risposte errate ma plausibili, definite anche "allucinazioni" o "confabulazioni" nel contesto dell'AI generativa, il mantenimento di una persona all'interno del processo rimane fondamentale. La tecnologia,

pur essendo potente, non è infallibile e la supervisione umana è essenziale per convalidare e contestualizzare le intuizioni generate dall'AI, soprattutto in scenari decisionali critici in ambito sanitario. La GAI ha applicazioni promettenti nell'assistenza sanitaria, ma i potenziali problemi etici devono essere affrontati in modo proattivo per evitare danni⁴⁹. Sono necessarie regole chiare e confini legali per allocare correttamente la responsabilità e proteggere gli utenti. Le preoccupazioni etiche derivano dalla potenziale interruzione del rapporto medico-paziente, dall'assistenza umana e da questioni di integrità. L'eccessivo affidamento all'intelligenza artificiale può minare la compassione ed erodere la fiducia.

In conclusione, l'introduzione della GAI nel settore sanitario rappresenta una svolta epocale, in particolare per quanto riguarda la gestione dei dati non strutturati. Le applicazioni della GAI nelle operazioni cliniche e nel processo decisionale promettono di migliorare l'efficienza e l'assistenza personalizzata dei pazienti. Tuttavia, le preoccupazioni sulla sicurezza dei dati, le distorsioni dei dati di addestramento e le considerazioni etiche sottolineano la necessità di un'implementazione attenta. Trovare un equilibrio tra innovazione e standard etici è fondamentale per realizzare il pieno potenziale della GAI nel plasmare il futuro dell'assistenza sanitaria.

Energia

L'industria energetica si trova di fronte a un triplice dilemma di sfide, che richiede soluzioni che garantiscano la sicurezza energetica, l'accessibilità economica e la transizione verso fonti energetiche più pulite. In risposta, l'Agenzia Internazionale dell'Energia (AIE) sottolinea il potenziale di trasformazione della digitalizzazione, in particolare attraverso la GAI, come catalizzatore per la riduzione dei costi, il miglioramento dell'efficienza e la riduzione delle emissioni nell'industria energetica. Mentre il settore è alle prese con questa trasformazione digitale, la complessità dei sistemi di alimentazione aumenta in modo esponenziale. I sistemi energetici stanno subendo trasformazioni significative, guidate dalla crescente domanda di elettricità e dall'intensificazione degli sforzi per la decarbonizzazione. A differenza delle strutture di rete

tradizionali, che indirizzavano l'energia dalle centrali elettriche centralizzate, i moderni sistemi energetici devono ora accogliere flussi multidirezionali di elettricità tra i generatori distribuiti, la rete e gli utenti finali. La proliferazione di dispositivi connessi alla rete, dalle stazioni di ricarica dei veicoli elettrici agli impianti solari residenziali, aggiunge un livello di imprevedibilità ai flussi di energia. Inoltre, la crescente integrazione tra il sistema elettrico e vari settori, tra cui i trasporti, l'industria e gli edifici, richiede un aumento sostanziale dello scambio di informazioni e strumenti più sofisticati per pianificare e gestire sistemi elettrici in evoluzione.

Questa complessità sottolinea il ruolo vitale della GAI nel promuovere l'efficienza e l'innovazione all'interno dell'industria energetica. I sistemi di alimentazione stanno diventando sempre più intensivi dal punto di vista dei dati, con le reti intelligenti che producono grandi quantità di dati. I contatori intelligenti, ad esempio, generano diverse migliaia di punti dati in più rispetto alle loro controparti analogiche, mentre i nuovi dispositivi per il monitoraggio dei flussi di energia della rete elettrica superano in

modo significativo la produzione di dati dei loro predecessori. Si stima che la sola flotta globale di turbine eoliche produca oltre 400 miliardi di punti dati all'anno. Le aziende energetiche riconoscono la GAI come una risorsa fondamentale che serve già per oltre 50 scopi diversi all'interno del sistema energetico. Si prevede che il mercato della tecnologia GAI nel settore raggiungerà i 13 miliardi di dollari, a testimonianza della sua crescente importanza nell'affrontare le sfide poste dall'evoluzione del panorama energetico.⁵⁰

L'intelligenza artificiale ha un potenziale di gran lunga superiore per accelerare la transizione energetica globale. Il World Economic Forum ha identificato nove "principi di AI per la transizione energetica" rivolti all'industria energetica, agli sviluppatori di tecnologie e ai responsabili politici. Se adottati, questi principi accelererebbero l'adozione di soluzioni di AI al servizio della transizione energetica, creando una comprensione comune di ciò che è necessario per sbloccare il potenziale dell'AI e di come adottare l'AI in modo sicuro e responsabile nel settore energetico. Tra le varie applicazioni che la GAI può avere nel processo di

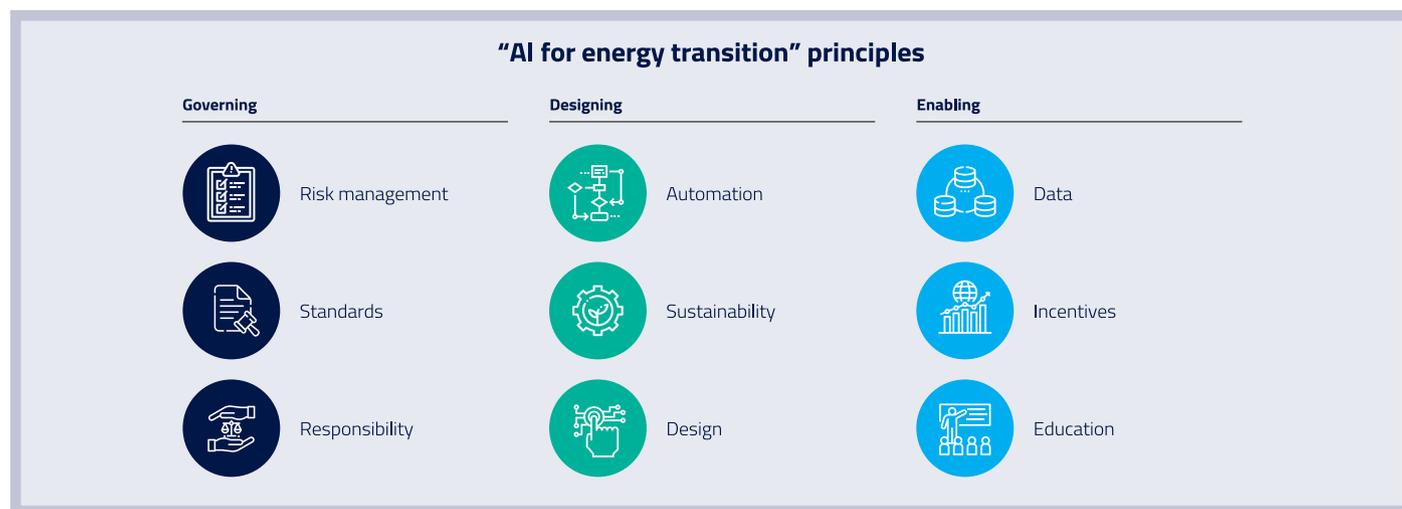


Figura 8 AI generativa per la transizione energetica, Fonte: Forum economico mondiale

transizione energetica, di seguito forniamo due esempi concreti per capire come la GAI può avere un impatto sul settore energetico.

1. Ottimizzazione energetica. La prima applicazione rilevante per la GAI è l'ottimizzazione energetica, un processo cruciale volto a migliorare l'efficienza e a ridurre gli sprechi di energia in diversi settori. Tradizionalmente, per ottimizzare l'energia si ricorre a soluzioni tecniche come l'aggiornamento delle apparecchiature, l'installazione di sensori e l'integrazione di fonti rinnovabili. Tuttavia, questi approcci devono affrontare sfide come gli elevati costi operativi, la limitata scalabilità, i problemi di compatibilità e gli ostacoli normativi alla partecipazione ai mercati energetici. La GAI emerge come forza trasformativa nell'affrontare queste sfide grazie alle sue capacità avanzate. Attraverso un'analisi completa dei dati, la GAI combina le informazioni provenienti da varie fonti, tra cui contatori intelligenti, previsioni meteo e preferenze degli utenti. Sfruttando le sue intuizioni analitiche, la GAI offre raccomandazioni attuabili per ridurre i costi energetici, migliorare l'efficienza complessiva dell'utilizzo dell'energia e implementare strategie per un'impronta energetica più sostenibile riducendo le emissioni di carbonio. Inoltre, la GAI consente di creare un ecosistema energetico dinamico, facilitando il commercio di energia peer-to-peer e la partecipazione attiva ai programmi di risposta alla domanda, contribuendo a un panorama energetico più resiliente e reattivo.

2. Reti intelligenti e risposta alla domanda. La seconda applicazione della GAI è la trasformazione e il potenziamento delle reti intelligenti. Guidato da pionieri del settore, questo approccio innovativo si concentra sulla creazione di modelli generativi guidati dall'intelligenza artificiale per i dati contenenti informazioni dei clienti. L'applicazione di questi modelli generativi si estende alla modellazione della rete e agli algoritmi di addestramento per le start-up dell'industria energetica. Grazie all'addestramento su dati esistenti, questi modelli generano dati realistici, consentendo alle parti interessate di esplorare scenari "what-if" al di là delle limitazioni degli attuali dataset. Ad esempio, i dati generati possono

prevedere il carico potenziale sulla rete con l'adozione di tecnologie solari da parte di un numero specifico di famiglie e come questo carico potrebbe fluttuare nel corso della giornata. L'integrazione di modelli di intelligenza artificiale generativa alle tecnologie delle reti intelligenti offre un campo di prova virtuale, consentendo lo sviluppo e il collaudo di prodotti hardware e software per la scalabilità e l'interoperabilità. Questo uso trasformativo dell'AI generativa promette di creare un futuro più sostenibile e resiliente per l'industria energetica.

Abbiamo riportato solo due esempi, ma il futuro dell'AI generativa nel settore dell'energia sembra promettente. L'intelligenza artificiale ha il potenziale per trasformare l'industria energetica, rendendola più efficiente, affidabile e sostenibile. Con la continua evoluzione della tecnologia, possiamo aspettarci algoritmi di AI più avanzati, una migliore analisi dei dati e una maggiore integrazione con altre promettenti tecnologie digitali come l'hardware neuromorfico e l'informatica quantistica. L'intelligenza artificiale ha il potenziale per accelerare la transizione verso un'economia a basse emissioni di carbonio, consentendo lo sviluppo di nuove tecnologie per le energie rinnovabili e migliorando l'efficienza di quelle esistenti.

In conclusione, l'industria energetica si trova ad affrontare sfide che richiedono soluzioni per la sicurezza, l'economicità e il passaggio a fonti più pulite. L'Agenzia Internazionale dell'Energia riconosce la GAI come catalizzatore trasformativo per la riduzione dei costi, il miglioramento dell'efficienza e la riduzione delle emissioni. Con l'evoluzione della complessità dei sistemi energetici, la GAI diventa fondamentale per l'efficienza e l'innovazione. La natura ad alta intensità di dati delle reti intelligenti e il crescente mercato della GAI nell'industria energetica ne evidenziano l'importanza. Due applicazioni chiave sottolineano il potenziale della GAI nell'ottimizzazione dei processi. In prospettiva, la GAI promette di migliorare l'efficienza, l'affidabilità e la sostenibilità, facilitando la transizione verso un'economia a basse emissioni di carbonio grazie ad algoritmi avanzati e all'integrazione alle tecnologie emergenti.

DAI RISCHI ALLA REGOLAMENTAZIONE E ALLE PROSPETTIVE DELL'UE

In questo capitolo iniziamo ad esaminare alcuni dei rischi posti dalla GAI che, in qualche modo, costituiscono la ragione alla base delle iniziative normative in corso, analizzate a livello globale (Sezione 4.2) e più in dettaglio a livello europeo (Sezione 4.3).

RISCHI POSTI DALL'AI GENERATIVA

La GAI promette di portare a compimento trasformazioni in vari ambiti, rimodellando radicalmente molti aspetti della nostra vita quotidiana. Tuttavia, l'avvento di questa nuova tecnologia comporta anche alcuni potenziali rischi. La letteratura affronta un'ampia serie di rischi che vanno dalle violazioni della proprietà intellettuale alla sicurezza informatica, e molti altri. Di seguito ci concentriamo selettivamente solo su alcuni rischi, in particolare su quelli che hanno effetti socioeconomici e politici più chiari e potenzialmente più consistenti.

Sebbene il dibattito sui rischi posti dall'AI non sia nuovo, i rapidi progressi nell'ambito della GAI richiedono una nuova prospettiva da parte di ricercatori e politici. Il dibattito sui rischi associati alla GAI si è ampliato particolarmente dopo la pubblicazione di un innovativo documento da parte di ricercatori dell'AI nel marzo 2021, intitolato "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" [N.d.T. I pericoli dei pappagalli stocastici: i modelli linguistici possono essere troppo grandi?]⁵¹. L'articolo ha causato la perdita del posto di lavoro in Google di due dei coautori e una successiva protesta da parte dei dipendenti di Google⁵². Secondo gli autori, che hanno coniato l'espressione "pappagalli stocastici", i modelli linguistici di grandi dimensioni (LLM) non fanno altro che ripetere o assemblare frasi basate su probabilità e modelli statistici appresi da vasti insiemi di testo, senza una reale comprensione o consapevolezza. Il documento

è il primo tentativo di considerare i rischi dei modelli linguistici di grandi dimensioni per quanto riguarda i loro costi ambientali e finanziari, la loro imperscrutabilità che comporta pregiudizi pericolosi sconosciuti, l'incapacità dei modelli di comprendere i concetti alla base di ciò che apprendono e la possibilità di usarli per ingannare le persone. Negli ultimi due anni sono state effettuate analisi molto esaustive dei rischi che emergono dalla GAI⁵³. Attingendo alla letteratura multidisciplinare dell'informatica, della linguistica e delle scienze sociali, possiamo identificare e strutturare il panorama dei rischi associati a questa tecnologia⁵⁴.

La prima area di rischio riguarda la potenziale discriminazione sociale. I modelli di GAI possono introdurre pregiudizi algoritmici dovuti a dati di addestramento imperfetti o a decisioni prese durante lo sviluppo del modello. Questo rischio rientra nei pericoli specifici⁵⁵. I LLM sono potenzialmente in grado di perpetuare e diffondere discriminazioni ingiuste, poiché i pregiudizi presenti nei dati di addestramento possono essere appresi e replicati dal modello. Spesso le risorse linguistiche su cui vengono addestrati i modelli linguistici dell'AI riflettono dati relativi a gruppi sociali dominanti, il che può sollevare problemi di diversità e inclusione. Questi rischi derivano in gran parte dalla selezione di dataset di addestramento che incorporano un linguaggio distorto e rappresentano in modo sproporzionato specifiche identità sociali. In alcuni casi, i dati di addestramento riflettono modelli storici di ingiustizia sistemica, quando sono raccolti da contesti in cui la disuguaglianza è lo status quo. In altri casi, i dati di addestramento possono essere distorti perché alcune comunità sono meglio rappresentate in tali dati rispetto ad altre, generando contenuti che non rappresentano la lingua di coloro che sono emarginati, esclusi o registrati con minore frequenza. Vi sono

diversi esempi concreti tratti dalla letteratura. Ad esempio, sono stati riscontrati pregiudizi di genere e di rappresentazione nelle storie di fantasia generate da GPT-3, dove i nomi femminili erano più spesso associati a storie sulla famiglia e sull'aspetto e descritti come meno potenti rispetto ai personaggi maschili⁵⁶. Pertanto, per ridurre al minimo le distorsioni dei modelli linguistici dell'AI, è essenziale utilizzare dataset di addestramento equilibrati che rappresentino in modo più equo i diversi gruppi.

La seconda area di rischio riguarda la diffusione di informazioni false o fuorvianti da parte dei LLM. I modelli linguistici e altri strumenti di GAI possono consentire la produzione di fake news, deepfakes e altre forme di contenuti manipolati che possono rivelarsi impossibili da distinguere da quelli reali. Nemmeno i modelli linguistici avanzati su larga scala sono in grado di prevedere in modo affidabile le informazioni reali, infatti emettono informazioni dettagliate e corrette in alcune circostanze, ma forniscono informazioni errate in altre. Ciò può accadere perché i LLM prevedono la probabilità di diverse parole successive in base alle parole precedenti. Tuttavia, il fatto che una frase sia probabile non indica in modo affidabile se la frase sia anche corretta dal punto di vista fattuale. Di conseguenza, non sorprende che i LLM assegnino frequentemente alte probabilità a previsioni false o insensate. I LLM che forniscono spesso informazioni corrette possono indurre gli utenti a fidarsi eccessivamente delle previsioni del modello, aggravando così i rischi che gli utenti si affidino a questi modelli quando sono inaffidabili o non sicuri. Probabilmente, un LLM che fornisce previsioni materialmente corrette il 99% delle volte può rappresentare un pericolo maggiore di uno che fornisce previsioni corrette il 50% delle volte, poiché è più probabile che le persone sviluppino una forte fiducia nel primo. Pertanto, la combinazione di AI e disinformazione può portare a credenze erronee su vasta scala che gli approcci tradizionali, come il fact-checking, l'educazione degli utenti e l'alfabetizzazione mediatica o gli strumenti di rilevamento, non sarebbero in grado di affrontare. Ciò rappresenta una minaccia significativa per l'equità dei processi democratici, in quanto può portare alla proliferazione della disinformazione, potenzialmente in grado di influenzare l'opinione pubblica e il processo decisionale.

Come discusso nella Sezione 3.1, l'impatto economico e organizzativo è un rischio sfaccettato nell'ambito del quale l'AI generativa può colpire la forza lavoro e avere un impatto negativo su gruppi specifici e comunità locali. Tali rischi sono difficili da prevedere, in parte a causa dell'incertezza sulla potenziale portata, sulla tempistica e sulla complessità di implementazione delle tecnologie linguistiche nell'economia. Gli effetti complessivi sull'occupazione dipenderanno anche dalla domanda di mansioni non automatizzate che continuano a richiedere l'impiego di personale umano, oltre che dalle più ampie tendenze macroeconomiche, industriali e commerciali. Tuttavia, come già accennato, l'automazione basata su LLM può avere conseguenze sulla qualità del lavoro e compromettere alcuni aspetti dell'economia creativa.

La privacy è un'altra considerazione importante in relazione alla GAI. Se le informazioni inserite dall'utente diventano identificabili negli output del modello, possono sorgere problemi relativi alla privacy, e ciò solleva preoccupazioni sulla divulgazione involontaria di dati sensibili. Possono verificarsi violazioni della privacy anche in presenza di processi di inferenza senza che i dati privati dell'individuo siano presenti nel dataset di addestramento. Analogamente ad altri modelli statistici, la GAI può realizzare inferenze corrette su una persona basandosi esclusivamente su dati correlazionali relativi ad altre persone, senza avere accesso a informazioni che potrebbero essere private su quel particolare individuo. Le previsioni di dati sensibili possono richiedere anche solo informazioni personali minime, come ad esempio chi un utente "segue" su Twitter. Non è chiaro se sia possibile trovare soluzioni semplici per mitigare i rischi di un'informazione senza introdurre nuove forme di censura o rendere inaccessibili le informazioni utili. Tuttavia, il potenziale rischio per la privacy derivante dai modelli di GAI può intaccare i diritti individuali, soprattutto se i dati sensibili sulle persone interessano la sfera pubblica e i processi democratici, come ad esempio nel caso in cui vengano diffuse informazioni personali sensibili durante le campagne elettorali.

LA REGOLAMENTAZIONE: UN'ANALISI GLOBALE SELETTIVA

Negli ultimi anni, in tutto il mondo sono stati pubblicati diversi quadri di governance dell'AI con l'obiettivo di fornire indicazioni di alto livello per uno sviluppo dell'AI sicuro e affidabile. Diverse organizzazioni multilaterali hanno pubblicato i propri principi, come i "Principi sull'intelligenza artificiale" dell'OCSE (2019), gli "Orientamenti etici per un'AI affidabile" dell'UE (2019) e le "Raccomandazioni sull'etica dell'intelligenza artificiale" dell'UNESCO (2021). Lo sviluppo dell'AI generativa, tuttavia, ha portato a nuove indicazioni, tra cui il "Processo di Hiroshima del G7 sull'Intelligenza Artificiale Generativa" (2023), recentemente pubblicato dall'OCSE. Questi principi e quadri di riferimento volontari fungono spesso da guida per i responsabili politici e di regolamentazione di tutto il mondo. A partire dal 2023, più di 60 paesi nelle Americhe, in Africa, in Asia e in Europa hanno pubblicato strategie nazionali sull'AI.⁵⁷

Sebbene nessun Paese abbia ancora approvato una normativa generale sull'AI o una normativa specifica sulla GAI, tra le principali iniziative legislative si annoverano quelle di Brasile, Cina, Unione Europea, India, Singapore, Corea del Sud e Stati Uniti. Gli approcci

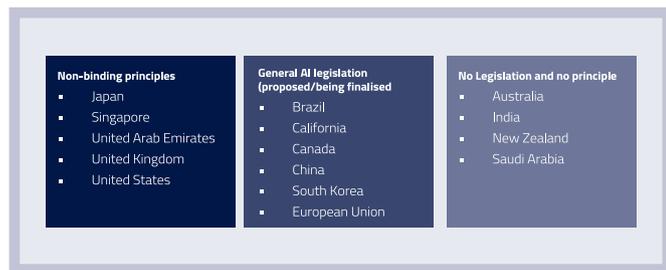


Figura 9 Normativa relativa all'AI nel mondo, dicembre 2023, selezione, Fonte: nostra elaborazione da Kremer et al (2023, cfr. nota 60);OECD.AI Policy observatory <https://oecd.ai/en/> ; <https://www.information-age.com/how-generative-ai-regulation-shaping-up-around-world-123503911/> ; <https://www.insideprivacy.com/artificial-intelligence/from-washington-to-brussels-a-comparative-look-at-the-biden-administrations-executive-order-and-the-eus-ai-act/>

adottati dai diversi Paesi variano da un'ampia regolamentazione dell'AI supportata da normative esistenti in materia di protezione dei dati e di sicurezza informatica (Unione Europea e Corea del Sud) a leggi specifiche per il settore (Stati Uniti) e approcci più generali basati su principi o linee guida (Brasile, Singapore e Stati Uniti). Ogni approccio presenta vantaggi e svantaggi e alcuni mercati passeranno nel tempo da linee guida basate su principi a una legislazione rigorosa⁵⁸. La figura esemplificativa che segue fornisce uno sguardo d'insieme.

Sebbene gli approcci varino, nel panorama normativo a livello globale sono emersi temi comuni:

- **Trasparenza.** L'obiettivo è la tracciabilità e la chiarezza dei risultati dell'AI e che gli utenti siano informati quando utilizzano un sistema di AI.
- **Intervento e supervisione umana.** I sistemi di AI dovrebbero essere sviluppati e utilizzati come strumenti al servizio delle persone, che rispettino la dignità umana e l'autonomia personale e che funzionino in modo da poter essere adeguatamente controllati e supervisionati dagli esseri umani.
- **Responsabilità.** Garantire la consapevolezza circa la responsabilità e i potenziali ricorsi relativi ai sistemi di AI.
- **Solidità tecnica e sicurezza.** Ridurre al minimo i danni involontari e imprevisti garantendo che i sistemi di AI siano solidi, ossia funzionino come previsto, rimangano stabili e possano correggere gli errori dell'utente.
- **Diversità, non discriminazione ed equità.** Un altro obiettivo è quello di garantire che i sistemi di AI siano privi di pregiudizi e che i risultati non comportino discriminazioni o trattamenti ingiusti delle persone.
- **Privacy e governance dei dati.** Lo sviluppo e l'utilizzo dei sistemi di AI devono rispettare le norme vigenti in materia di privacy e protezione dei dati, elaborando dati che soddisfino standard elevati di qualità e integrità.
- **Benessere sociale e ambientale.** L'AI deve essere sostenibile per le persone e per l'ambiente.

Di seguito esaminiamo, con diverso grado di dettaglio, l'iniziativa in alcuni Paesi (o enti amministrati, come nel caso della California).

Partiamo dagli Stati Uniti e facciamo un confronto con l'Europa⁵⁹. La differenza principale tra l'approccio dell'UE e quello degli Stati Uniti sta nel fatto che, mentre l'AI Act introduce una nuova e sistematica regolamentazione dell'AI, l'Ordine esecutivo firmato dal Presidente Biden il 30 ottobre 2023 (si veda più avanti) non crea nuovi obblighi legislativi e introduce piuttosto diverse indicazioni per le agenzie governative. Negli Stati Uniti, negli ultimi anni sono emersi diversi documenti di orientamento e quadri volontari, come l'"AI Risk Management Framework" del National Institute of Standards and Technology (NIST) statunitense, una guida volontaria pubblicata nel gennaio 2023, e il "Blueprint for an AI Bill of Rights" della Casa Bianca, una serie di principi di alto livello pubblicati nell'ottobre 2022. Tuttavia, le due principali proposte di legge a livello federale sono l'"Algorithmic Accountability Act" e l'"AI Disclosure Act", entrambe in fase di discussione. La prima mira a creare tutele per gli individui soggetti a processi decisionali algoritmici in settori come l'alloggio, il credito, l'istruzione e altro ancora. La legge incarica la Federal Trade Commission di creare regolamenti che forniscano alle aziende che gestiscono questi sistemi di intelligenza artificiale istruzioni concrete di valutazione e procedure per una valutazione continua. Inoltre, l'AI Disclosure Act richiederebbe l'inclusione, in ogni risultato generato dall'AI, del seguente avviso: "Disclaimer: questo risultato è stato generato da un'intelligenza artificiale" Più recentemente, il 30 ottobre 2023, il Presidente Biden ha emanato l'"Executive Order Safe, Secure, and Trustworthy Artificial Intelligence"⁶⁰. Pur differenziandosi dall'approccio dell'AI Act dell'UE, l'ordine esecutivo statunitense condivide l'attenzione per i "modelli di fondazione a duplice uso", in particolare quelli con capacità ad alte prestazioni in compiti che comportano rischi per la sicurezza, l'economia o la salute pubblica. Questi rischi comprendono aree come le minacce informatiche, le armi chimiche/biologiche/radiologiche/nucleari e l'inganno/manipolazione. L'ordine esecutivo statunitense introduce misure di "red-teaming" per gli sviluppatori di modelli di fondazione di AI per valutare i potenziali rischi e le vulnerabilità, particolarmente importanti per le applicazioni a duplice uso. Gli sviluppatori sono tenuti a comunicare al governo i risultati del red-teaming, specificando come affrontano i rischi e le vulnerabilità più significative nei loro sistemi.

La posizione degli Stati Uniti conferma che i modelli di fondazione richiedono un approccio distinto dalle altre applicazioni di AI.

La strategia cinese nei confronti dell'AI, invece, è caratterizzata da un duplice obiettivo: promuovere l'innovazione in ambito AI e mantenere la supervisione statale su questa tecnologia. Questo impegno è evidente in documenti politici chiave, come il Piano di sviluppo dell'intelligenza artificiale di nuova generazione, che delinea l'ambizione della Cina di porsi come leader nel panorama globale dell'AI entro il 2030. Contrariamente alla legge europea in materia, che adotta un approccio globale e orizzontale con standard adattabili a diverse applicazioni di AI, la Cina impiega un metodo più mirato e sistematico, utilizzando leggi distinte per affrontare sfide specifiche dell'AI. Attualmente, le normative cinesi sull'AI si sono concentrate su due questioni distinte: gli algoritmi di raccomandazione basati sull'AI e gli strumenti di deep synthesis (sintesi profonda), comunemente impiegati nella creazione di deepfakes. Per quanto riguarda gli algoritmi di raccomandazione basati sull'AI, le normative impongono ai fornitori di servizi di questo settore di limitare la discriminazione, di ridurre la diffusione di informazioni negative e di affrontare la questione delle condizioni di sfruttamento lavorativo dei fattorini. Questa legislazione conferisce inoltre ai consumatori cinesi il diritto di disattivare le raccomandazioni algoritmiche e di ricevere spiegazioni quando gli algoritmi influenzano in modo significativo i loro interessi. Nel campo della deep synthesis, le norme cinesi comprendono l'uso di algoritmi per generare o modificare artificialmente i contenuti online. Tale legislazione prevede che i contenuti di deep synthesis aderiscano ai controlli sulle informazioni e siano chiaramente etichettati come generati dalla deep synthesis, e impone ai fornitori di implementare misure che ne impediscano l'uso improprio. Inoltre, include requisiti di censura vagamente definiti e obbliga i fornitori a registrare i propri algoritmi. Se da un lato la Cina impiega ampiamente l'AI nelle sue forze dell'ordine e nell'apparato di sorveglianza, dall'altro sono stati introdotti regolamenti per disciplinare l'uso della tecnologia da parte di enti non governativi. Secondo queste norme, gli strumenti di riconoscimento facciale sono consentiti solo per scopi specifici,

quando gli strumenti alternativi sono inadeguati, e il loro impiego negli spazi pubblici deve avere come priorità la sicurezza pubblica.

In Australia, il bilancio federale del 2023 ha visto l'annuncio di una Responsible AI Network [rete per l'AI responsabile] da parte del governo, oltre a un finanziamento di 41,2 milioni di dollari australiani (26,9 milioni di dollari) per l'introduzione responsabile di diverse tecnologie di AI in tutto il Paese. Inoltre, sono in corso colloqui tra le autorità di regolamentazione in merito a possibili modifiche alla legge sulla privacy più generale, al fine di coprire la mancanza di trasparenza che può verificarsi con l'addestramento di modelli di AI tramite cicli di feedback senza la presenza di una supervisione umana. Si discute anche sull'uso di dati analitici e biometrici per addestrare i modelli, che potrebbero richiedere l'introduzione di ulteriori norme sulla privacy.

In Brasile il processo legislativo in materia di AI è stato anticipato da un quadro giuridico approvato dal governo nel settembre 2022, ma ampiamente criticato perché troppo vago. Il processo legislativo in materia di AI in Brasile è stato anticipato da un quadro giuridico approvato dal governo nel settembre 2022, ma ampiamente criticato perché troppo vago. Dopo l'uscita di ChatGPT nel novembre 2022, il dibattito tra i lobbisti ha condotto all'invio di un report al governo brasiliano contenente raccomandazioni su come regolamentare l'intelligenza artificiale. Il documento è attualmente in fase di discussione all'interno del governo brasiliano e le date per le prossime stesure non sono ancora state annunciate.

La regolamentazione dell'innovazione in materia di AI in California, dove si trova il cuore della Silicon Valley, è destinata a essere una sfida sempre attuale per le autorità di regolamentazione, vista la presenza di OpenAI e dei principali investitori Microsoft e Google, che hanno la loro sede nello Stato e ne supervisionano il forte coinvolgimento. Per superare questa sfida complessa, le autorità di regolamentazione stanno pianificando una proposta ad ampio respiro. Basandosi sul quadro nazionale dato dal AI Bill of Rights, la legislazione su cui si sta lavorando mira a prevenire la discriminazione e i danni in tutti i settori privati, tra cui l'istruzione,

i servizi pubblici, l'assistenza sanitaria e i servizi finanziari. Come misure di salvaguardia sono state suggerite valutazioni d'impatto annuali che gli sviluppatori e gli utenti dovrebbero presentare al Dipartimento per i diritti civili della California e che descriverebbero in dettaglio i tipi di strumenti automatizzati coinvolti, e l'accesso pubblico. Inoltre, si prevede di chiedere agli sviluppatori di implementare un quadro di governance che descriva in dettaglio le modalità di utilizzo della tecnologia e le possibili conseguenze.

In Canada è prevista una legge sull'intelligenza artificiale e i dati (Artificial Intelligence and Data Act, AIDA) che sarà adottata non prima del 2025, la cui stesura è stata avviata nell'ambito del Bill C-27, il Digital Charter Implementation Act, 2022. In tal modo, il governo canadese si allinea a normative simili negli Stati Uniti e nell'UE, con l'intenzione di basarsi sulla legge canadese esistente in materia di protezione dei consumatori e dei diritti umani per riconoscere la necessità che i sistemi di AI "ad alto impatto" rispettino le leggi sui diritti umani e sulla sicurezza. Inoltre, si afferma che il Ministro dell'Innovazione, della Scienza e dell'Industria avrebbe il compito di garantire che la regolamentazione stia al passo con l'evoluzione tecnologica e che si possano creare nuove disposizioni di legge contro l'uso doloso. Sono state individuate sei aree di obblighi principali a cui i sistemi ad alto impatto devono attenersi: responsabilità, equità e correttezza, analisi e monitoraggio umano, sicurezza, trasparenza, validità e solidità.

In India il governo ha annunciato nel marzo 2021 che avrebbe dato un "lieve cambio" alla regolamentazione dell'AI con l'obiettivo di preservare l'innovazione in tutto il Paese, senza piani immediati per una regolamentazione specifica. Optando per un approccio contrario alla regolamentazione dello sviluppo dell'AI, il Ministero dell'Elettronica e dell'Informatica ha riconosciuto questo settore tecnologico come "significativo e strategico", ma ha dichiarato che metterà in atto politiche e misure infrastrutturali per contribuire alla lotta contro pregiudizi, discriminazioni e questioni etiche. Il governo indiano ha proposto dei quadri volontari per la gestione dell'AI. La sua strategia nazionale per l'intelligenza artificiale del 2018 ha preso in considerazione cinque settori chiave di sviluppo dell'AI: agricoltura,

istruzione, sanità, città intelligenti e mobilità intelligente. Nel 2020, poi, gli usi etici dell'AI sono stati descritti in dettaglio in una bozza della strategia nazionale per l'intelligenza artificiale, che chiede che tutti i sistemi siano trasparenti, responsabili e imparziali. In netto cambiamento rispetto agli sviluppi precedenti, il 1° marzo 2024 il governo indiano ha emesso un avviso in cui si chiede alle piattaforme di richiedere l'autorizzazione esplicita del Ministero dell'Elettronica e della Tecnologia dell'Informazione prima di distribuire qualsiasi "modello di Intelligenza Artificiale/LLM/AI generativa, software o algoritmo inaffidabile" per "gli utenti di Internet in India"⁶¹. Inoltre, chiede agli intermediari o alle piattaforme di garantire che i propri sistemi non consentano pregiudizi o discriminazioni né minino l'integrità del processo elettorale nonché di etichettare tutti i contenuti multimediali e i testi creati tramite AI attraverso identificatori o metadati unici, così che siano facilmente identificabili.

In Corea del Sud la legge sull'AI è attualmente alle fasi finali di elaborazione, in attesa di votazione all'interno dell'Assemblea nazionale. La legge, nella sua forma attuale, intende chiarire che la normativa deve consentire a qualsiasi utente di creare nuovi modelli senza dover ottenere alcuna pre-approvazione da parte del governo, mentre i sistemi considerati "ad alto rischio" per la vita dei cittadini devono ottenere una fiducia a lungo termine. La proposta di legge in esame si concentra sull'innovazione nazionale in un'ottica etica prevedendo che le imprese che utilizzano l'AI generativa ricevano un supporto governativo su come sviluppare i sistemi in modo responsabile. Inoltre, la Commissione per la protezione delle informazioni personali del Paese ha annunciato l'intenzione di creare una task force dedicata al ripensamento della protezione dei dati biometrici, alla luce degli sviluppi dell'AI generativa.

Allo stato attuale, la regolamentazione dell'AI generativa nel Regno Unito è destinata a rimanere nelle mani delle autorità di regolamentazione del settore in cui l'AI viene utilizzata, senza che sia prevista una legge generale oltre al GDPR britannico. Negli annunci ufficiali su questo tema, il governo ha optato per un "approccio favorevole all'innovazione", dimostrando che il paese

vuole assumere la leadership nella corsa globale all'intelligenza artificiale. Tuttavia, rimangono ancora i rischi dell'AI generativa come le violazioni di sistema, la disinformazione e i pregiudizi. Per contribuire a mitigare questo problema, il governo britannico ha pubblicato una valutazione d'impatto al fine di determinare una regolamentazione adeguata ed equa per gli sviluppatori di AI. Questa misura fa parte della più ampia Strategia nazionale per l'AI, la cui sintesi recita: «Esiste una serie di fallimenti del mercato (asimmetria informativa, incentivi disallineati, esternalità negative, carenze normative), il che significa che i rischi dell'AI non vengono affrontati in modo adeguato». «Il governo britannico è nella posizione migliore per proporre un regime normativo intersettoriale adatto a raggiungere questi obiettivi». Tra gli obiettivi fissati vi sono la promozione della crescita delle PMI nel settore AI, l'aumento della fiducia del pubblico e il mantenimento o il miglioramento della posizione del Regno Unito nello Stanford Global AI Index. L'Autorità garante della concorrenza e dei mercati (CMA), nel frattempo, ha avviato un'analisi dei modelli di base dell'AI, esaminando lo sviluppo di strumenti quali ChatGPT per considerazioni relative alla concorrenza e alla tutela dei consumatori. Gli sviluppatori di AI sono chiamati a dimostrare l'allineamento con cinque principi generali: sicurezza, protezione e solidità; trasparenza e spiegabilità adeguate; equità; responsabilità e governance; contestabilità e ricorso.

Nel 2023, i tentativi internazionali di governare e salvaguardare l'AI hanno guadagnato slancio grazie a una serie di iniziative multistakeholder di grande impatto, plasmando un approccio globale unificato per affrontare le sfide più urgenti. L'organo consultivo sull'intelligenza artificiale del Segretario Generale delle Nazioni Unite, co-presieduto da Carme Artigas e James Manyika e convocato dal Segretario Generale António Guterres, è pronto a pubblicare il proprio report finale prima dell'atteso Summit del Futuro nell'estate del 2024, fornendo preziose indicazioni per uno sviluppo responsabile dell'intelligenza artificiale⁶². Contemporaneamente, il Partenariato Globale sull'Intelligenza Artificiale (Global Partnership on Artificial Intelligence, GPAI) ha tenuto il proprio Summit 2023

a Nuova Delhi, promuovendo la collaborazione sulle priorità legate all'AI in diversi settori⁶³. L'accordo dei leader del G7 sui Principi Guida Internazionali sull'Intelligenza Artificiale, insieme a un Codice di Condotta volontario per gli sviluppatori di AI nell'ambito del processo di Hiroshima sull'AI, ha segnato un passo significativo nella governance globale dell'AI. Gli undici principi guida forniscono una guida completa per le organizzazioni che sviluppano e distribuiscono sistemi avanzati di AI, con l'obiettivo di promuovere la sicurezza, l'affidabilità e l'uso responsabile. Il Processo di Hiroshima sull'Intelligenza Artificiale del G7, istituito nel maggio 2023, contribuisce a un più ampio discorso internazionale sulle misure di contenzione per l'AI, allineandosi con le discussioni in seno all'OCSE, al GPAI e nel contesto del Consiglio per il Commercio e la Tecnologia UE-USA e dei Partenariati Digitali dell'UE⁶⁴. Inoltre, l'AI Safety Summit 2023 a Bletchley Park e il REAIM 2023 all'Aia, nei Paesi Bassi, sono serviti come importanti piattaforme internazionali per la collaborazione, affrontando i rischi e le opportunità associate all'AI. Queste iniziative collettive sottolineano l'impegno del mondo a muoversi nello sviluppo e nella diffusione responsabile delle tecnologie dell'AI su scala globale.

Il 5 settembre 2024, il Consiglio d'Europa un'organizzazione che riunisce anche Paesi non facenti parte della Ue ma collocati sul territorio europeo (per un totale di 46 Stati membri), ha approvato la Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law. La Convenzione quadro, sottoscritta dalla Commissione Europea che ha negoziato anche a nome della Santa Sede, degli Stati Uniti, del Canada, del Messico, del Giappone, di Israele, dell'Australia, dell'Argentina, del Perù, dell'Uruguay e del Costa Rica, è la prima convenzione internazionale in materia di intelligenza artificiale. Il testo dedica poca attenzione alle specificità delle AI generativa, ma include regole che mirano a garantire che l'intelligenza artificiale non comprometta i valori fondamentali dell'Unione Europea, come il rispetto della dignità umana e i diritti fondamentali.

REGOLAMENTAZIONE: L'AI ACT DELL'UE

L'AI Act dell'UE, proposto dalla Commissione europea nell'aprile 2021, non contemplava esplicitamente la GAI. Dopo il fenomenale decollo di ChatGPT nel novembre 2022, il processo legislativo e i negoziati sull'AI Act sono dovuti ripartire da capo e la GAI e i modelli di fondazione sono diventati il punto chiave della controversia tra le istituzioni europee (Commissione europea, Parlamento europeo e Consiglio, il cosiddetto Trilogo). Il 9 dicembre 2023, i responsabili politici dell'UE hanno raggiunto un accordo politico sulla legge sull'AI, anche se, come vedremo, c'è ancora molto da fare e l'AI Act non sarà pienamente attuato fino al 2026.

Quando ChatGPT ha avuto un'impennata nel 2022, i circoli politici di Bruxelles hanno iniziato a preoccuparsi del fatto che la struttura della legge sull'AI riflettesse in gran parte la tradizionale legislazione dell'UE sulla responsabilità del prodotto, inadeguata per la GAI. A differenza dei prodotti stabili con finalità specifiche, la GAI può essere impiegata per diversi scopi non previsti durante lo sviluppo; ciò ha posto la GAI al di fuori dell'ambito di applicazione della legge sull'AI, così come è stata redatta, fino alla fine del 2022. Un problema correlato risiede nell'approccio della Commissione Europea basato sul rischio, che presuppone che i sistemi di AI possano essere confinati in determinate classi di rischio. La parte centrale della legge sull'AI nella proposta iniziale della Commissione⁶⁵ è la regolamentazione dei sistemi di AI ad alto rischio. I sistemi di AI ad alto rischio possono far parte di un componente di sicurezza di un prodotto, oppure sono un prodotto stesso, o ancora sono applicazioni specifiche in un settore designato come ad alto rischio. Tuttavia, la GAI possiede la capacità di svolgere diversi compiti e di apprendere di nuovi senza un uso predeterminato. Pertanto, non può essere considerata ad alto rischio, il che implica che gli sviluppatori di AI diventerebbero responsabili della conformità solo se mettessero a punto o adattassero in modo significativo il sistema di AI per un uso ad alto rischio. Un altro problema è che gli attori che sviluppano i modelli di fondazione delle applicazioni di GAI trarrebbero profitto da un'applicazione downstream lontana, evitando di assumersi le relative responsabilità, anche se ci sono diversi esempi che

dimostrano che i modelli spesso comportano rischi intrinseci integrati nei dati sottostanti. Consapevole di questi problemi, nell'aprile del 2023 l'organizzazione di ricerca AI Now Institute ha pubblicato un report firmato da oltre 50 esperti e istituzioni in cui si sostiene che i sistemi di AI di uso generale «comportano gravi rischi e non devono essere esentati dalla prossima legge europea sull'AI». Il report sostiene che lo sviluppo dei modelli di fondazione comporta dei rischi, come le potenziali violazioni della privacy commesse per raccogliere i dati necessari all'addestramento di un modello, che possono essere affrontati solo regolamentando i modelli stessi piuttosto che la loro applicazione.⁶⁶

Alla luce di queste sfide, sia il Consiglio europeo che il Parlamento europeo hanno avanzato proposte specifiche tra aprile e luglio 2023 per adattare l'AI Act al fine di includere l'AI generativa e i relativi modelli di fondazione. Nella sua posizione⁶⁷ il Consiglio ha introdotto la nuova categoria dei sistemi di AI per scopi generali, definiti come segue: «per sistema di AI per scopi generali si intende un sistema di AI che, indipendentemente dal modo in cui viene immesso sul mercato o messo in servizio, anche come software open-source, è destinato dal fornitore a svolgere funzioni generalmente applicabili [...] un sistema di AI per scopi generali può essere utilizzato in una pluralità di contesti ed essere integrato in una pluralità di altri sistemi di AI»;⁶⁸ Il Consiglio ha proposto che le norme sui sistemi ad alto rischio si applichino all'AI per scopi generali. La posizione del Parlamento⁶⁹ si è invece concentrata sui modelli di fondazione e ha introdotto una categoria di rischio separata. Ha definito i modelli di fondazione come «un modello di AI addestrato su un'ampia gamma di dati su vasta scala, [che] è progettato per la generalità dell'output e può essere adattato a un'ampia gamma di compiti specifici»⁷⁰. Gli emendamenti proposti dal Parlamento richiedono che tutti i fornitori di modelli di fondazione svolgano una due diligence di base sulle loro offerte⁷¹ e soddisfino tre obblighi: a) identificazione del rischio. Anche se non è possibile identificare in anticipo tutti i potenziali casi di utilizzo di un modello di base, i fornitori sono in genere consapevoli di alcuni vettori di rischio. OpenAI

sapeva, ad esempio, che il dataset di addestramento per GPT-4 presentava alcune distorsioni linguistiche, poiché oltre il 60% di tutti i siti web sono in inglese. Il Parlamento europeo renderebbe obbligatorio identificare e mitigare i rischi ragionevolmente prevedibili, in questo caso l'imprecisione e la discriminazione, con il supporto di esperti indipendenti; b) Test. I fornitori sono tenuti a fare scelte progettuali adeguate per garantire che il modello di base raggiunga livelli appropriati di prestazioni, prevedibilità, interpretabilità, correggibilità, sicurezza e cybersecurity. Poiché il modello di base funge da elemento costitutivo per molti sistemi di AI downstream, esso deve soddisfare alcuni standard minimi. Ad esempio, le aziende il cui prodotto di AI si basi su PaLM 2 devono essere certe che il building block soddisfi i requisiti di cybersecurity fondamentali; c) documentazione. I fornitori di modelli di fondazione dovranno fornire una documentazione sostanziale sotto forma di schede tecniche, schede modello e istruzioni d'uso comprensibili. Questo è essenziale non solo per aiutare i fornitori di sistemi di AI downstream a capire meglio cosa stanno perfezionando o mettendo a punto, ma anche per consentire loro di rispettare eventuali requisiti normativi.

L'accordo provvisorio sull'AI Act, raggiunto dai negoziatori dell'UE il 9 dicembre 2023 e approvato dal Parlamento europeo il 13 marzo 2024, ha ripreso alcune delle proposte avanzate dal Parlamento.⁷² Tale accordo è stato precedentemente osteggiato da Francia, Germania e Italia che hanno spinto per un regime normativo più leggero per i modelli che supportano i sistemi di intelligenza artificiale per scopi generali come ChatGPT e Bard. Questi tre Paesi hanno chiesto di limitare le regole in questo settore ai codici di condotta, per non ostacolare le start-up europee come Mistral AI e Aleph Alpha che potrebbero sfidare le aziende americane. Tuttavia, il Parlamento europeo è stato unito nel chiedere regole severe per questi modelli. Il compromesso si è basato su un approccio graduale, con regole di trasparenza orizzontali per tutti i modelli e obblighi aggiuntivi per i modelli obbligatori che si ritiene comportino un rischio sistemico. Il compromesso si riferisce al termine di sistemi/modelli di AI per scopi generali ("GPAL")

e distingue tra obblighi su due livelli: (1) una serie di obblighi orizzontali che si applicano a tutti i modelli GPAI e (2) una serie di obblighi aggiuntivi per i modelli GPAI con rischio sistemico. Per quanto riguarda il primo livello, tutti i fornitori di modelli GPAI dovranno aderire ai requisiti di trasparenza redigendo, tra l'altro, una documentazione tecnica. Dovranno inoltre rispettare la legge sul copyright dell'UE e fornire riepiloghi dettagliati sui contenuti utilizzati per la formazione. I modelli GPAI di livello inferiore saranno esentati dai requisiti di trasparenza finché sono in fase di ricerca e sviluppo o se sono open source. Per quanto riguarda il secondo livello, i modelli GPAI saranno indicati come comprensivi del rischio sistemico quando soddisfano determinati criteri. I modelli GPAI che sono stati classificati come di rischio sistemico saranno soggetti a obblighi più stringenti, che comprendono «la conduzione di valutazioni dei modelli, la valutazione e la mitigazione dei rischi sistemici, la conduzione di test avversariali, la segnalazione alla Commissione di incidenti gravi, la garanzia di sicurezza informatica e la rendicontazione sulla loro efficienza energetica». I modelli GPAI con rischio sistemico possono essere conformi all'AI Act dell'UE aderendo a codici di condotta, almeno fino a quando non saranno pubblicati standard UE armonizzati. I codici di condotta saranno sviluppati dall'industria, dalla comunità scientifica, dalla società civile e da altre parti interessate insieme alla Commissione. Inoltre, un gruppo scientifico di esperti indipendenti emetterà avvisi sui rischi sistemici e supporterà la classificazione e la verifica dei modelli. Questo compromesso è chiaramente meno severo nei confronti delle GPAI rispetto alla proposta originale del PE.

L'accordo raggiunto il 9 dicembre, tuttavia, non era ancora concluso al 100% e richiedeva un ulteriore lavoro tecnico. La complessità della legge ha fatto sì che il suo perfezionamento tecnico richiedesse più di un mese. Questo lavoro tecnico prevede il perfezionamento di 95 premesse soprattutto dell'art. 5 che riguarda quei sistemi che in linea di principio dovrebbero essere vietati, ma che saranno esentati dal divieto in quanto necessari per alcuni particolari interessi pubblici. Il 24 gennaio, la presidenza belga del Consiglio dei ministri dell'UE ha presentato la versione finale del testo. In quell'occasione, la maggior

parte degli Stati membri ha mantenuto delle riserve, non avendo avuto il tempo sufficiente per analizzare il testo in modo completo. Tali riserve sono state infine sciolte con l'adozione dell'AI Act da parte del Comitato dei Rappresentanti Permanenti il 2 febbraio 2024, quando gli ambasciatori dell'UE hanno votato all'unanimità l'adozione della legge confermando l'accordo politico raggiunto nel dicembre 2023⁷³. Successivamente, il 13 febbraio 2024, la nuova versione dell'AI Act è stata approvata dalle commissioni per il mercato interno e per le libertà civili del Parlamento europeo. Il 13 marzo 2024 il testo è stato approvato in maniera definitiva dal Parlamento Europeo. A tale votazione ha fatto seguito l'approvazione da parte del Consiglio della Ue, avvenuta il 21 maggio 2024. Infine, con il Regolamento (UE) 2024/1689 del 13 giugno 2024 è giunto a compimento il percorso legislativo dell'AI Act. Il Regolamento è stato pubblicato in Gazzetta Ufficiale Europea il 12 luglio 2024, ed è entrato in vigore il 2 agosto dello stesso anno. I divieti diventeranno applicabili sei mesi dopo l'entrata in vigore della legge e le disposizioni per i "modelli di fondazione" (GAI) dopo 12 mesi. Il periodo di transizione per i sistemi inclusi nell'Allegato 2 durerà fino al 2026. Mentre alcune disposizioni dell'AI Act si applicheranno già subito dopo l'adozione del regolamento, altre (ad esempio, alcuni requisiti sui sistemi di AI ad alto rischio) saranno applicabili solo al termine di un periodo transitorio. In questo contesto, la Commissione sta avviando il Patto sull'AI (AI Pact)⁷⁴, cercando l'impegno volontario dell'industria ad anticipare l'AI Act e ad iniziare ad attuarne i requisiti prima dei termini legali. Dopo l'entrata in vigore della legge sull'AI, c'è comunque ancora molto da fare, tra cui la creazione dell'Ufficio AI della Commissione, l'assunzione di esperti, la stesura di linee guida e soprattutto la formulazione di standard orizzontali e verticali entro il 2025. Tali standard sono di importanza cruciale perché devono trasformare le affermazioni generali in affermazioni tecniche che possono essere misurate e verificate. Ad esempio, frasi come "adeguate misure di gestione del rischio" dovranno essere rese operative e concrete. La legge è piena di affermazioni generali e vaghe che devono essere supportate da standard tecnici chiari. Pur essendo considerati "tecnici", una volta formulati possono avere un forte impatto sulle implicazioni legislative della legge. Molte cose possono ancora accadere durante la definizione

di tali standard e l'organizzazione della supervisione normativa, e molta acqua sarà passata sotto i ponti entro il 2026, quando la legge sarà pienamente attuata. Il panorama tecnologico da qui al 2026 potrebbe cambiare radicalmente. Quindi, da qui al 2026 la legge sull'AI rimane un bersaglio mobile ancora circondato da incertezze su come sarà implementata operativamente e su quali saranno i suoi effetti e la sua applicabilità dato il ritmo del cambiamento tecnologico.

Le principali fonti di insoddisfazione del Parlamento europeo (PE) nei confronti dell'ultima versione dell'accordo raggiunto sono principalmente tre⁷⁵. In primo luogo, fin dall'inizio il PE si è espresso a favore di una legge generale sui diritti e di linee guida etiche applicabili a tutti i sistemi di AI (senza l'approccio della stratificazione del rischio), da rendere poi operativa a livello verticale per la sanità, i media, le industrie creative, ecc. e da integrare con una nuova legge sulla responsabilità. Il Parlamento europeo non era e non è soddisfatto dell'approccio orizzontale scelto dalla Commissione, in quanto solleverebbe diverse ambiguità e problemi a causa della mancanza di un contesto verticale specifico. Questo problema potrebbe essere risolto solo introducendo standard tecnici orizzontali e verticali molto solidi e chiari. In secondo luogo, il PE avrebbe preferito un approccio molto più centralizzato all'applicazione delle norme. Nella soluzione attuale, le questioni riguardanti l'AI in relazione alla legge dovranno essere sottoposte prima alle autorità nazionali competenti, che a loro volta si coordineranno semplicemente con l'Ufficio AI della Commissione. L'Ufficio AI funzionerà fondamentalmente come organo di coordinamento e come piattaforma per la condivisione delle buone pratiche. Questa soluzione, secondo il PE, potrebbe portare alla stessa mancanza di attuazione unificata in tutta l'UE che caratterizza il GDPR, e lascerebbe autonomia agli Stati membri in cui hanno sede i giganti delle Big Tech di essere indulgenti con loro. In terzo luogo, il PE ritiene che le dichiarazioni e gli articoli sui "modelli di fondazione" (GAI) siano troppo vaghi e deboli, a causa delle pressioni esercitate all'ultimo minuto dalle aziende sostenute da Francia e Germania. In particolare, secondo il PE, la procedura di designazione (come GPAI con rischi sistemici) descritta nell'articolo 8 lascia delle scappatoie alle grandi aziende tecnologiche per contestare le decisioni di designazione

e portarle in tribunale. Infine, il PE disapprova l'esenzione prevista per i sistemi open-source per quanto riguarda la trasparenza e la documentazione, in quanto è controintuitiva e contraria allo spirito dell'open-source. Vale la pena evidenziare che tra le aziende europee che operano nello spazio open-source ci sono la francese Mistral e la tedesca Aleph Alpha, che hanno entrambe criticato la legge, come riportiamo di seguito. È inoltre importante ricordare che in seguito Mistral ha avviato una partnership con Microsoft. Così, Mistral si è presentata come campione europeo e ha fatto pressioni per ridurre al minimo la regolamentazione della GAI nella legge per consentire alle start-up europee di competere con i giganti tecnologici statunitensi per poi collaborare con uno di loro. Molti politici dell'UE, con il senno di poi, pensano di essere stati presi in giro e che Mistral abbia fatto da copertura per Microsoft.

Secondo Metakides⁷⁶, il lavoro sull'AI Act è stato completato nonostante le forti pressioni esercitate dalle lobby, che sono arrivate all'ultimo momento (20 ottobre 2023) soprattutto contro la regolamentazione dei "modelli di fondazione" (GAI). Le big tech hanno utilizzato non solo il budget autodichiarato per l'attività di lobbying, ma anche le PMI, le associazioni di consumatori, i think tank, le ONG, le società di consulenza per fare pressione contro le norme severe in generale e soprattutto contro la regolamentazione della GAI. Secondo Metakides, il Consiglio rimane la scatola nera della politica dell'UE. È la meno trasparente tra le istituzioni dell'UE per quanto riguarda l'attività di lobbying. È difficile per i cittadini sapere chi esercita pressioni sui rappresentanti del loro governo in seno al Consiglio. Questa volta le aziende europee hanno fatto pressione insieme alle Big Tech statunitensi contro la regolamentazione della GAI. La già citata società francese Mistral, il cui responsabile delle relazioni con l'UE è Cedric O (ex segretario di Stato per il digitale), ha lanciato una campagna di lettera aperta firmata da oltre 150 PMI in cui si afferma che: «La legge sull'intelligenza artificiale metterebbe a rischio la competitività e la sovranità tecnologica dell'UE», dichiarazione ripresa poi dal Ministero francese. Una posizione simile è stata assunta dalla tedesca Aleph Alpha, a cui ha fatto eco il Ministero tedesco.

Dopo aver fornito un resoconto del processo passato, della situazione attuale e dei possibili sviluppi futuri della legge sull'AI, riportiamo di seguito, in modo molto selettivo, alcune opinioni e posizioni tra le tante che sono state espresse negli ultimi dodici mesi.

La prima opinione tradizionale è quella proposta da personaggi come Anu Bradford. Secondo Bradford, l'UE ha sfidato l'argomentazione spesso sostenuta che l'AI non può essere regolamentata e che è troppo presto per intervenire data la natura in rapida evoluzione di questa tecnologia. Al contrario, sostiene Bradford, l'AI Act rappresenta un esempio potente e molti governi lo stanno già esaminando. Si potrà forse osservare un effetto Bruxelles, poiché è probabile che diversi sviluppatori di AI si conformino all'AI Act anche al di fuori dell'Europa. Dopotutto, non è detto che vogliano addestrare modelli diversi per ogni singolo mercato in cui operano⁷⁷. Si tratta quindi dell'«Effetto Bruxelles» versione 3.0 (dopo il GDPR e il DMA).

D'altra parte, una seconda posizione è quella degli autori che sostengono che l'imposizione di tutti gli obblighi in modo uniforme a tutti i fornitori di modelli di fondazione, indipendentemente dalle dimensioni, potrebbe soffocare l'innovazione e rafforzare la posizione dominante sul mercato di aziende leader come OpenAI, Anthropic e Google Deepmind⁷⁸. Per le aziende più piccole può essere difficile competere e raggiungere questi leader di mercato. Per risolvere questo problema, diverse parti interessate hanno proposto una distinzione tra modelli di fondazione di importanza sistemica e modelli di fondazione di importanza non sistemica, con oneri notevolmente più leggeri per questi ultimi. Questa proposta si allineerebbe all'approccio della Legge sui servizi digitali dell'UE, riconoscendo l'importanza di adattare gli obblighi di due diligence in base al tipo, alle dimensioni e alla natura del servizio. La distinzione tra modelli di fondazione sistemici e non sistemici, e l'imposizione di obblighi completi solo ai primi, sarebbe giustificata dalle maggiori risorse di cui dispongono le imprese che sviluppano modelli sistemici, rendendole meglio attrezzate per la conformità alla normativa. Inoltre, è probabile che il danno potenziale causato da una deviazione dalla conformità da parte di una piccola impresa con pochi clienti sia

significativamente inferiore a quello causato da un modello di base sistemico. Un approccio di questo tipo richiederebbe alcuni criteri specifici per distinguere i modelli di fondazione «sistemici» o «ad alto impatto» dagli altri. Gli studiosi hanno accennato a criteri che potrebbero essere utilizzati per identificare i diversi tipi di modello, come le fonti di dati utilizzate o le risorse informatiche necessarie per l'addestramento iniziale del modello⁷⁹.

Una terza posizione esprime un netto scetticismo sulla necessità di regolamentare orizzontalmente soprattutto le GAI. Secondo Soete⁸⁰, ad esempio, la regolamentazione dell'AI è resa difficile dalla velocità di sviluppo delle piattaforme e degli strumenti commerciali di AI e le sandbox regolamentari assomigliano a «sabbie mobili». Soete sostiene che le autorità di regolamentazione dovrebbero invece concentrarsi sull'apertura della «scatola nera dei dati utilizzati per addestrare l'AI e supportare gli algoritmi», assicurandosi che i dati siano resi pubblici e trasparenti, «in modo da sfruttare il potenziale per un'applicazione molto più ampia con ulteriori miglioramenti, sulla base di prove ed errori»⁸¹. Cita l'approccio proposto da Mons⁸² di rendere gli input di dati «Fully AI Ready» (ovvero «pienamente pronti per l'AI») in tutte le future ricerche sull'AI (un gioco di parole alternativo all'originale FAIR (giusto), acronimo di Findable, Accessible, Interoperable and Reusable, ossia reperibile, accessibile, interoperabile e riutilizzabile, per le macchine così come per le persone⁸³). La posizione di Mons è che «i tentativi di regolamentare questi strumenti, e il concomitante entusiasmo, possono solo favorire gli interessi commerciali dei loro creatori»⁸⁴. Suggestisce che l'entusiasmo verso questo modello potrebbe implodere su se stesso e che, piuttosto che una regolamentazione dall'alto verso il basso, ci si dovrebbe concentrare su una buona gestione dei dati. Conclude che «cercare un approccio dall'alto verso il basso per limitare tali risultati spingerebbe l'apprendimento automatico in aree in cui le normative non possono arrivare. Un input affidabile, la continua esposizione e la critica nei confronti di tale entusiasmo sono fondamentali, così come i principi di dare l'esempio e di alimentare i modelli con substrati e vincoli concettuali adeguati»⁸⁵.

PROSPETTIVE DELL'UE SULL'AI GENERATIVA

A prescindere dal dibattito normativo e dal modo in cui la legge sull'AI prenderà forma una volta definiti gli standard tecnici, la dura realtà è che l'Europa è attualmente in ritardo in materia di AI rispetto alla concorrenza globale e alla leadership internazionale. In questo passaggio epocale, l'Europa manca di un'iniziativa di AI coesa e ambiziosa a livello europeo e fatica ad affermarsi con prodotti o infrastrutture di AI "Made in Europe". L'attuale panorama globale dell'AI accentua la vulnerabilità dell'Europa: solo il 6% dei finanziamenti per l'AI è destinato alle start-up dell'UE-27, a fronte del 61% assegnato alle aziende statunitensi e del 17% a quelle cinesi. In particolare, le aziende legate all'AI generativa con sede negli Stati Uniti si sono assicurate ben 23,8 miliardi di dollari dal 2019 al 2023 (di cui 12,3 miliardi di dollari solo per OpenAI), che costituiscono il 75% degli investimenti totali in tali aziende durante tale periodo. Le domande di brevetto hanno un andamento simile: il 34% proviene da inventori statunitensi, il 22% da inventori cinesi e solo l'11% da inventori dell'UE27. Ciononostante, l'Europa possiede importanti talenti nel campo dell'AI, eccellenza scientifica e impegno verso uno sviluppo responsabile dell'AI, offrendo le potenzialità per tracciare un percorso verso un'AI affidabile.

In linea con la Figura precedente, una stima recente mostra che il 73% dei modelli di fondazione dell'AI proviene dagli Stati Uniti, dove lo sviluppo è guidato principalmente dalle grandi aziende tecnologiche, e il 15% dalla Cina⁸⁶. Le start-up europee devono affrontare sfide come la raccolta di fondi, l'accesso alle competenze necessarie, l'utilizzo di dati in lingue meno diffuse e l'ottenimento di un'infrastruttura informatica sufficiente. Esistono tuttavia alcune start-up e iniziative europee degne di nota, come le francesi LightOn e Mistral, la tedesca Aleph Alpha, che stanno

sviluppando LLM alternativi che salvaguardano i dati e incorporano i valori europei. Tuttavia, come già anticipato all'inizio della sezione 3.2, Mistral, pur essendo stata presentata in precedenza come un modello europeo, ha recentemente avviato una partnership con Microsoft, che ha investito 16 milioni di euro nell'azienda francese e che presto renderà disponibili i modelli di AI dell'azienda attraverso la sua piattaforma di cloud computing Azure⁸⁷. Allo stesso modo, Open GPT-X, un'iniziativa di dieci organizzazioni tedesche del mondo degli affari, della scienza e dei media, sta sviluppando la risposta europea a GPT-3. Il Ministero federale tedesco per gli Affari economici e l'Azione per il clima finanzia il progetto Open GPT-X nell'ambito dell'iniziativa di finanziamento Gaia-X con circa 15 milioni di euro. Sotto la guida degli Istituti Fraunhofer per l'Analisi Intelligente e i Sistemi Informativi (AII) e per i Circuiti Integrati (IIS), il progetto OpenGPT-X si prefigge di sviluppare un modello di linguaggio dell'AI di grandi dimensioni per l'Europa, che offra protezione dei dati e diversità linguistica europea. Tuttavia, se paragonati agli 11 miliardi di dollari raccolti da OpenAI, questi modelli di LLM emergenti in Europa dispongono di una quantità limitata di finanziamenti. Un'altra iniziativa europea degna di nota è Poro, un modello linguistico aperto sviluppato in Finlandia da Silo AI, il più grande laboratorio privato di intelligenza artificiale in Europa, insieme all'Università di Turku e all'HPLT. Si tratta di un passo importante per il rafforzamento della sovranità digitale europea e per la democratizzazione dell'accesso ai modelli linguistici di grandi dimensioni (LLM) per tutte le lingue europee. Il modello è la prova del successo dell'applicazione di un metodo innovativo per l'addestramento di LLM per lingue con scarse risorse.

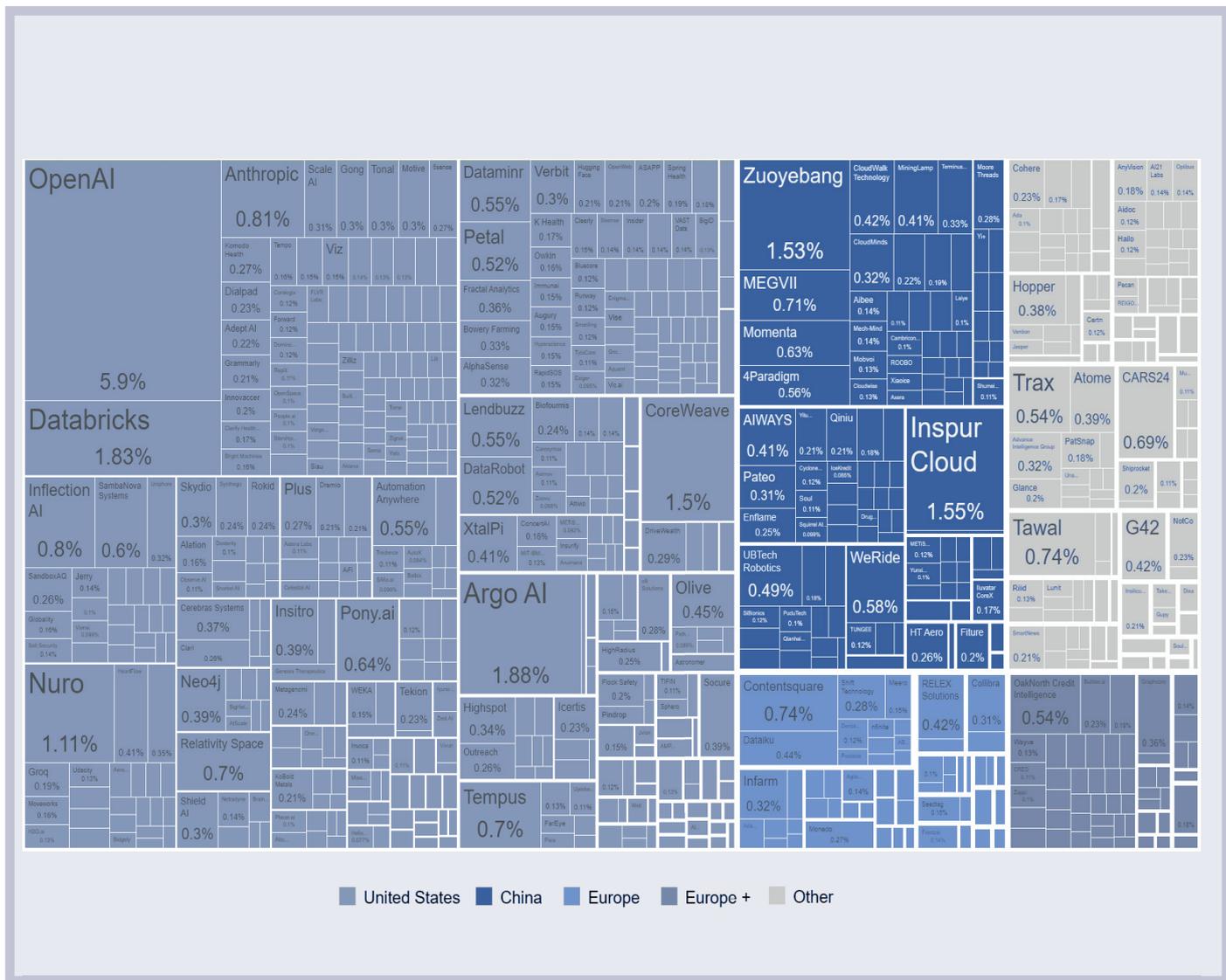


Figura 10 La quota di finanziamenti di venture capital (VC) ricevuti dalle start-up dell'AI, Fonte: Crunchbase

L'esame dei brevetti degli ultimi otto anni rivela un panorama concentrato sull'innovazione dell'AI, con oltre il 25% di tutte le domande di brevetto europee sull'AI provenienti da quattro regioni specifiche: il centro di Londra, l'Île-de-France (Parigi), il Noord-Brabant (Eindhoven) e l'Ober Bayern (Monaco). Questa concentrazione mette in evidenza il talento e il potenziale innovativo dell'AI in alcune città e regioni europee selezionate. Tale concentrazione suggerirebbe investimenti sostanziali nella ricerca e nello sviluppo dell'AI, con un approccio mirato a sostenere e promuovere questi poli di innovazione, in modo che possano diventare magneti globali per i finanziamenti e i talenti. Inoltre, l'Europa manca attualmente di coesione interregionale, poiché la maggior parte delle connessioni tra brevetti di AI in Europa rimane circoscritta all'interno dei confini nazionali. Per soddisfare la domanda globale di sofisticate tecnologie di AI, l'Europa dovrebbe dare priorità e potenziare la connettività inter-hub per un ecosistema di AI collaborativo e interconnesso⁸⁸.

Lo sviluppo dell'ecosistema europeo dell'AI generativa si collega alla discussione sulla legge sull'AI e sul modo in cui la nuova normativa potrebbe avere un impatto sulle aziende europee emergenti nel settore dell'AI. Una ricerca condotta da un gruppo di associazioni europee di AI alla fine del 2022 ha rilevato che il 50% delle start-up intervistate riteneva che la legge sull'AI avrebbe rallentato l'innovazione dell'AI in Europa, mentre il 16% ha dichiarato che stava considerando di interrompere lo sviluppo dell'AI o di trasferirsi al di fuori dell'UE⁸⁹. Più di recente, i governi francese, tedesco e italiano si sono espressi a favore di un approccio alla regolamentazione dell'AI "equilibrato e favorevole all'innovazione", che sia basato sui rischi ma che riduca anche "gli inutili oneri amministrativi per le imprese che ostacolerebbero la capacità dell'Europa di innovare". I governi francese e tedesco hanno dimostrato attivamente il loro impegno a promuovere l'innovazione nei rispettivi settori dell'AI. A giugno, il presidente Macron si è impegnato a stanziare 500 milioni di euro per sostenere i "campioni" dell'AI, mentre il governo tedesco, ad agosto, ha annunciato l'intenzione di quasi raddoppiare i finanziamenti pubblici per la ricerca sull'AI, portandoli

a circa 1 miliardo di euro nei prossimi due anni. Entrambi i governi condividono la preoccupazione che un'eccessiva regolamentazione possa ostacolare la crescita dei rispettivi settori nazionali dell'AI e suggeriscono che la legge europea sull'AI debba concentrarsi sulla regolamentazione degli usi dell'AI piuttosto che sui modelli di fondazione sottostanti.

In fase di ultimazione del presente report, il 24 gennaio 2024 la Commissione europea ha pubblicato la comunicazione "sulla promozione delle start-up e dell'innovazione nell'intelligenza artificiale affidabile"⁹⁰, realizzando un pacchetto di innovazione sull'AI per sostenere le start-up e le PMI dell'intelligenza artificiale. Tra gli aspetti più degni di nota di questo pacchetto segnaliamo i seguenti:

- Una modifica del regolamento EuroHPC per istituire le fabbriche di AI, un nuovo pilastro per le attività dell'impresa comune dell'UE nel settore dei supercomputer. In questa rientrano:
 - L'acquisizione, l'aggiornamento e la gestione di supercomputer dedicati all'AI per consentire un machine learning rapido e il training di grandi modelli di AI per scopi generali (GPAI).
 - Facilitare l'accesso ai supercomputer dedicati all'AI, contribuendo ad ampliare l'uso dell'AI a un gran numero di utenti pubblici e privati, comprese le start-up e le PMI.
 - Offrire uno sportello unico per le start-up e gli innovatori, sostenendo le start-up dell'AI e l'ecosistema della ricerca nello sviluppo di algoritmi, nella valutazione dei test e nella convalida di modelli di AI su larga scala, mettendo a disposizione strutture di programmazione facili da usare per i supercomputer e altri servizi di abilitazione all'AI.
 - Consentire lo sviluppo di una serie di applicazioni emergenti di AI basate su modelli di AI per scopi generali.
- Sostegno finanziario della Commissione attraverso Horizon

Europe e il programma Digital Europe dedicato all'AI generativa. Questo pacchetto genererà un investimento complessivo aggiuntivo pubblico e privato di circa 4 miliardi di euro fino al 2027.

- Iniziative di accompagnamento per rafforzare il pool di talenti generativi dell'UE nel campo dell'AI attraverso attività di istruzione, formazione, qualificazione e riqualificazione.
- Incoraggiare ulteriormente gli investimenti pubblici e privati nelle start-up e nelle scale-up dell'AI, anche attraverso il capitale di rischio o il sostegno azionario (anche tramite le nuove iniziative del Programma EIC Accelerator e di InvestEU).
- L'accelerazione dello sviluppo e della diffusione di spazi comuni europei di dati messi a disposizione della comunità dell'AI, per la quale i dati sono una risorsa fondamentale per addestrare e migliorare i propri modelli. Oggi è stato pubblicato anche un nuovo documento di lavoro sugli spazi comuni europei di dati che fornisce lo stato di avanzamento dei lavori.
- Diversi Stati membri uniranno le forze attraverso il consorzio per un'infrastruttura digitale europea per l'alleanza per le tecnologie del linguaggio (European Digital Infrastructure Consortium for the Alliance for Language Technologies, ALT-EDIC), che fornirà un accesso centralizzato alle risorse linguistiche per lo sviluppo di LLM europei. Poiché i modelli avanzati possono gestire efficacemente più tipi di dati contemporaneamente (ad esempio testo, audio, video, immagini, codice), ALT-EDIC aprirà anche la possibilità di applicazioni di AI più olistiche e complete in vari domini.
- L'iniziativa "GenAI4EU", che mira a sostenere lo sviluppo di nuovi casi d'uso e di applicazioni emergenti nei 14 ecosistemi industriali europei e nel settore pubblico. Tra le aree di applicazione rientrano la robotica, la salute, le biotecnologie, la produzione, la mobilità, il clima e i mondi virtuali.

Questo pacchetto innovazione va considerato come la terza tappa di un quadro di governance dell'AI a livello europeo, la prima delle quali è la legge sull'AI e la seconda la politica di concorrenza e l'antitrust. La Comunicazione sollecita investimenti pubblici e si

inserisce nella strategia europea di messa a disposizione di tutti i tre ingredienti fondamentali della GAL: potenza informatica (rete di fabbriche di AI), dati (mercato unico) e talento (formazione). Sembra quindi che gli investimenti pubblici vadano di pari passo con la regolamentazione.

CONSIDERAZIONI COMPARATIVE CONCLUSIVE

Sebbene l'Europa sia la patria di alcune delle migliori università e dei migliori talenti del mondo, i progressi sono ostacolati dall'enfasi posta sulle strategie nazionali separate per l'AI, dalla lentezza del processo decisionale, dagli investimenti troppo limitati nell'AI da parte del settore privato europeo e dalle procedure burocratiche del settore pubblico. Man mano che l'adozione di questi modelli linguistici di grandi dimensioni (LLM) prende piede, l'Europa si trova di fronte alla prospettiva di una crescente dipendenza dai modelli di intelligenza artificiale stranieri, che rappresenta una potenziale minaccia per la competitività generale dell'economia europea. Ma non è in gioco solo l'economia. Come abbiamo discusso nelle sezioni precedenti, si prevede che l'AI generativa avrà un impatto importante sulla società e un'area a rischio è la cultura. Il dominio assoluto delle grandi aziende statunitensi e cinesi di AI è sconcertante. Non disponendo di piattaforme di AI europee valide, gli utenti europei dovranno accettare le imposizioni etico-culturali dell'AI sviluppata in altre parti del mondo e la diversità linguistico-culturale europea potrebbe risentirne. Come mostra la classifica sottostante, solo quattro Paesi dell'UE (Germania, Finlandia, Paesi Bassi e Francia) entrano nella top 10 della classifica dell'AI (vedi sotto).

L'attuale posizione dell'Europa non deve essere scambiata per una mancanza di talento o di potenziale tecnologico. In realtà, diversi hub europei ospitano un ricco bacino di talenti dell'AI, eccellenza scientifica e impegno per uno sviluppo responsabile dell'AI, che può forgiare un percorso verso un'AI affidabile, radicata nei valori umanitari e democratici. Se si considerano le città e le regioni, il quadro è molto meno allarmante rispetto a quello a livello nazionale. Le migliori start-up di intelligenza artificiale di

Parigi, ad esempio, ricevono finanziamenti simili a quelli di Boston e più di quelle con sede a Pittsburgh, Seattle, Chicago, Shenzhen, Guangzhou o Hangzhou. Ciò che manca in Europa è una “politica industriale orientata alla missione”, qualcosa di simile alla missione di Airbus di promuovere l’eccellenza europea nell’AI⁹¹.

	Overall	Talent	Infrastructure	Operating	Research	Development	Government Strategy	Commercial	Scale	Intensity
1	United States	1	1	28	1	1	8	1	1	5
2	China	20	2	3	2	2	3	2	2	21
3	Singapore	4	3	22	3	5	16	4	10	1
4	United Kingdom	5	24	40	5	8	10	5	4	10
5	Canada	6	23	8	7	11	5	7	7	7
6	South Korea	12	7	11	12	3	6	18	8	6
7	Israel	7	28	23	11	7	47	3	17	2
8	Germany	3	12	13	8	9	2	11	3	15
9	Switzerland	9	13	30	4	4	56	9	16	3
10	Finland	13	8	4	9	14	15	12	13	4
11	Netherlands	8	16	15	10	13	28	20	11	8
12	Japan	11	5	10	20	6	18	23	6	25
13	France	10	11	25	15	18	13	10	9	20

Figura 11 Classifica globale dell’Intelligenza Artificiale: i 10 migliori operatori, Fonte: <https://www.tortoisemedia.com/intelligence/global-ai/#rankings>

TECNOLOGIA, DINAMICHE SOCIOECONOMICHE E REGOLAMENTAZIONE: **GLI SCENARI**

PRINCIPALI AMBITI DI INCERTEZZA

Nei capitoli precedenti abbiamo presentato l'AI generativa come tecnologia (Capitolo 2), le sue dinamiche socioeconomiche (Capitolo 3), i principali rischi per la società e la democrazia e il processo normativo (Capitolo 4), seguiti da una valutazione dello stato di avanzamento dei lavori nell'UE (Capitolo 5). Lo sviluppo tecnologico, le dinamiche socioeconomiche, il processo normativo e il corrispondente livello di investimenti pubblici sono tutti fonte di incertezza.

La tecnologia alla base dell'AI generativa può essere vista come il principale motore del cambiamento con un impatto sia sulle dinamiche socioeconomiche che sulla regolamentazione. Nell'attuale panorama tecnologico l'attenzione si concentra su modelli generalisti, monolitici e closed source, che richiedono risorse su larga scala e concentrate in termini di investimenti, dati e potenza di calcolo. Tuttavia, vi sono incertezze riguardo al futuro sviluppo di questo panorama tecnologico, poiché in queste tre dimensioni esistono alternative: modelli generalisti vs. specialisti, modelli monolitici vs. federativi e approcci open-source vs. closed-source allo sviluppo e alla formazione dei modelli. Questi sviluppi tecnologici potrebbero aumentare o diminuire la concorrenza, incidendo sui modelli di business e sui flussi di reddito. La tecnologia potrebbe anche rendere la regolamentazione superflua o obsoleta o generare nuove opportunità associate a nuovi rischi che richiederebbero una nuova regolamentazione. Se lo sviluppo tecnologico ridurrà il bisogno di grandi quantità di dati per addestrare i modelli e diminuirà, di conseguenza, anche i requisiti di potenza di calcolo, saranno garantite basse barriere all'ingresso, una concorrenza vivace e minori opportunità di concorrenza sleale

da parte dei grandi operatori storici. Al contrario, se la tecnologia si svilupperà in modo tale da richiedere quantità sempre maggiori di dati e potenza di calcolo, ciò ostacolerà la concorrenza e incentiverà pratiche concorrenziali sleali di bundling che aumenteranno ulteriormente la concentrazione del mercato. Un'altra incertezza, in parte legata alla tecnologia e in parte all'attuazione della normativa esistente, riguarda la misura in cui i dati proprietari accumulati dagli operatori storici che gestiscono una piattaforma costituiranno un vantaggio insormontabile, o se gli operatori più piccoli saranno in grado di pre-addestrare e mettere a punto il proprio modello in modo efficace ed efficiente utilizzando il numero crescente di database accessibili al pubblico, e se, in Europa, in base alle disposizioni del DMA, agli operatori più piccoli sarà concesso l'accesso ai dati posseduti dai grandi operatori storici. Gli sviluppi tecnologici, abbinati all'autoregolamentazione, potrebbero anche rendere superflua la regolamentazione se i modelli futuri riducessero l'insorgere dei rischi sociali descritti nel paragrafo 4.1.

Le dinamiche socio-economiche dipendono dalle tendenze del mercato, dalle decisioni di investimento, dalle strategie competitive delle aziende, da come le autorità antitrust terranno sotto controllo i grandi operatori storici, dagli sviluppi del mercato del lavoro, nonché dalle azioni intraprese da consumatori e dei cittadini, sia in termini di adozione che di reclami e richieste di riequilibrio. A questo proposito, vale la pena notare che vi è una differenza negli atteggiamenti e nei comportamenti nel nostro diverso ruolo di "cittadini" e di "consumatori/utenti". Consideriamo, ad esempio, il caso delle grandi piattaforme online e dei servizi digitali. In Europa questi sono stati regolamentati (dal DMA e dal Digital Service Act,

DSA) per tutelare i nostri diritti di cittadini, eppure molti di noi che si comportano da consumatori/utenti continuano a utilizzarli con scarsa attenzione alle nuove disposizioni normative, per abitudine (pregiudizio dello status quo), lock-in (rete di contatti e capitale reputazionale) e convenienza. Lo stesso potrebbe accadere con le applicazioni e le soluzioni di AI generativa, se lo sviluppo tecnologico dovesse essere tale da generare convenienza e produrre un'ottima esperienza d'uso per gli utenti/consumatori, mentre coloro che esprimono preoccupazioni e critiche come cittadini rimarrebbero una minoranza attiva ma sempre più esigua, riducendo così il sostegno all'intervento normativo. La dinamica socioeconomica dipende anche dal successo o dal fallimento di settori strategici di applicazioni come, ad esempio, la sanità e l'energia. Non è certo se gli investimenti continueranno a concentrarsi principalmente sui produttori di modelli o se i finanziamenti si estenderanno anche ad altri livelli dell'ecosistema GAI, e soprattutto se tali investimenti continueranno a essere orientati verso gli Stati Uniti e la Cina o raggiungeranno anche l'UE. Inoltre, non si sa ancora con certezza se i ricavi si concentreranno a favore dei primi operatori di successo e degli altri operatori storici o se si distribuiranno in modo più uniforme. Anche il futuro panorama competitivo è molto incerto, con fattori che possono favorire la concentrazione oligopolistica e tendenze contrastanti in grado di mantenere condizioni di parità. Non è certo fino a che punto i grandi operatori storici adotteranno metodi di concorrenza sleale (bundling e tying) e le autorità antitrust saranno in grado di fermarli. L'adozione da parte di aziende e consumatori condiziona notevolmente l'ulteriore sviluppo della tecnologia e il suo impatto sociale ed economico, soprattutto sulla produttività e sulla crescita, ma anche a livello occupazionale. Infine, poiché lo sviluppo e l'adozione della tecnologia avanzano inevitabilmente, fintanto che la tecnologia non si perfezionerà e non sarà regolamentata, alcuni dei rischi descritti in precedenza si concretizzeranno e produrranno conseguenze, e resta da vedere in che misura i consumatori e i cittadini si mobilitano per difendere i propri diritti e ricevere un trattamento adeguato. Uno di questi rischi è molto imminente e riguarda le prossime elezioni europee del 2024, quando la legge sull'intelligenza artificiale non sarà

ancora pienamente applicabile e i sistemi GAI potrebbero essere utilizzati in modo malevolo per orchestrare fake news e campagne di disinformazione per influenzare i risultati delle elezioni.

La governance dell'AI, che comprende la regolamentazione (AI Act), gli investimenti pubblici (pacchetto innovazione), la politica della concorrenza e l'antitrust, è ancora molto incerta. In questa sede ci riferiamo all'AI Act sia in generale che con riferimento specifico alla GAI. Il lavoro non è ancora finito e la legge sarà pienamente attuata nel 2026. Nel 2025 dovranno essere prodotte norme e linee guida tecniche orizzontali e verticali. La disposizione relativa alla GAI attualmente contenuta nella legge rimane vaga. Ad esempio, la procedura per la designazione di un modello GPAI come di livello 1 (con meno obblighi) o di livello 2 (modelli GPAI con rischi sistemici soggetti a maggiori obblighi) è vaga e lascia alle imprese la possibilità di impugnarla in tribunale. Inoltre, i modelli GPAI con rischio sistemico possono conformarsi all'AI Act aderendo a codici di condotta, almeno fino a quando non saranno pubblicati standard UE armonizzati, probabilmente ma non certamente entro il 2025. Infine, i modelli GPAI di livello inferiore saranno esentati dai requisiti di trasparenza mentre sono in fase di R&S o se sono open source. È lecito concludere che la regolamentazione della GAI è tutt'altro che certa e chiara. Vale anche la pena di notare che, come anticipato, alcuni analisti sostengono che l'imposizione di tutti gli obblighi in modo uniforme a tutti i fornitori di modelli di fondazione, indipendentemente dalle dimensioni, potrebbe soffocare l'innovazione e rafforzare la posizione dominante sul mercato di aziende leader come OpenAI, Anthropic e Google Deepmind⁹². Al momento, quindi, non si può sapere fino a che punto la legge sarà applicata in modo rigido e rigoroso, in stretta collaborazione tra le autorità nazionali e la Commissione, con standard tecnici chiari, o se rimarrà in qualche misura poco rigorosa, gestita per lo più a livello nazionale, e richiederà molte interpretazioni aggiuntive ed eventualmente chiarimenti nei tribunali. Vale anche la pena di notare che l'AI Act non contiene alcuna disposizione in materia di concorrenza, e quindi la dimensione normativa non incorpora la politica di concorrenza e le iniziative antitrust. In questo contesto

È importante sottolineare il ruolo cruciale che la standardizzazione avrà nel rendere più concrete e operative molte delle disposizioni attualmente generali e generiche. Gli standard giocheranno un ruolo strategico nell'AI ACT e molto dipenderà da quali attori vi parteciperanno e dalla misura in cui le PMI europee riusciranno a influenzarli. Inoltre, l'entità e l'orientamento dei futuri investimenti pubblici rimangono incerti, in quanto il pacchetto innovazione presentato dalla Commissione nel gennaio 2024 per ora rimane un piano ancora da attuare e da impiegare. Infine, oltre alle iniziative normative come l'AI Act, ci sono incertezze riguardo alla politica sulla concorrenza e alle iniziative antitrust all'interno del quadro di governance della GAI. Esiste il rischio di pratiche di concorrenza sleale da parte degli operatori storici, come l'accorpamento di prodotti o l'assunzione di accordi di esclusiva, ed è ancora incerto in che misura tali pratiche saranno attuate dalle aziende e contrastate/mitigate dalle autorità antitrust. In funzione del modo in cui la politica sulla concorrenza e l'antitrust affronteranno tali questioni, potrebbero diminuire o aumentare le potenziali barriere all'ingresso per i nuovi operatori in termini di accesso ai dati, ai talenti e alla potenza di calcolo.

Queste tre dimensioni interagiscono evidentemente tra loro e al fine di rendere tali complesse interazioni reciproche sarebbe necessario un sofisticato esercizio di modellazione dinamica del sistema. Con qualche semplificazione concettuale, scegliamo le dinamiche socioeconomiche e la regolamentazione come le due dimensioni dell'incertezza che danno forma agli scenari e le osserviamo da una prospettiva di alto livello. La tecnologia sarà utilizzata per caratterizzare lo scenario individuato. Le incertezze tecnologiche sulla necessità futura di infrastrutture di calcolo e dati proprietari e la supremazia di architetture di modelli chiusi e monolitici rispetto a modelli più aperti e federati richiedono un monitoraggio continuo per anticipare le esigenze e le opportunità future. D'altra parte, non è certo se gli investimenti pubblici sosterranno o meno gli innovatori e le start-up dell'AI per compensare la posizione dominante degli operatori storici, e ciò influirà sulla struttura competitiva del mercato.

GLI SCENARI PROPOSTI

Alla luce della discussione precedente, è stato possibile selezionare le due incertezze che rappresentano gli assetti degli scenari rappresentati qui di seguito: "le dinamiche socioeconomiche" e "la regolamentazione". Come spiegato, è la tecnologia stessa il principale motore che avrà un impatto sia sul processo normativo che sulle dinamiche socio-economiche e che, pertanto, sarà inserita all'interno delle storyline degli scenari.

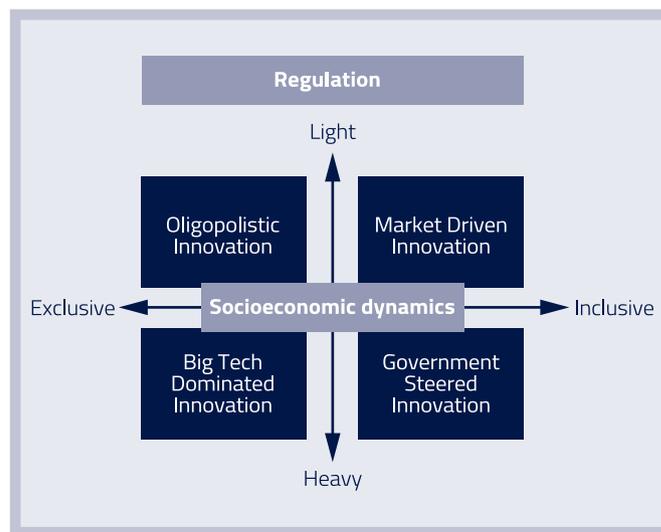


Figura 12 Gli scenari, Fonte: Elaborazione degli autori

Le dinamiche socioeconomiche variano tra due estremi considerati inclusivi sul lato destro ed esclusivi sul lato sinistro. Questa caratterizzazione delle dinamiche socioeconomiche è legata al livello di concorrenza in generale, ma anche al tipo di operatori che saranno in grado di rimanere sul mercato, nonché alla misura in cui la chiusura o l'apertura del mercato europeo saranno attuate. Così, ad esempio, una dinamica più inclusiva comporterebbe che le PMI europee siano in grado di competere e che le start-up europee riescano a ingrandirsi, mentre una dinamica esclusiva

comporterebbe il dominio degli operatori storici, sotto forma di big tech o di big tech più un numero limitato di nuove aziende che riescono a consolidarsi, con meno spazio per le PMI e le start-up dell'UE.

La normativa varia fra gli estremi di pesante e leggera, per riflettere le incertezze ancora presenti nell'AI Act, soprattutto per quanto riguarda la GAI, e il rapporto poco chiaro con la politica della concorrenza. Un regime normativo snello sarà anche flessibile e lascerà spazio all'innovazione degli operatori di mercato, ma comporterà anche incertezze durature per quanto riguarda l'attuazione e l'applicazione, che potrebbero richiedere un'interpretazione ed eventualmente una sentenza del tribunale. Un regime normativo severo comporta non solo controlli più rigorosi e costanti dei fornitori con vari adempimenti e sanzioni (presentazione di documentazione trasparente, audit, divieti e multe), ma anche una definizione più stringente e chiara dei livelli di rischio, degli standard, delle linee guida e delle procedure da applicare. Questo regime porterebbe certamente alla certezza del diritto, ma potrebbe anche soffocare l'innovazione delle start-up, delle PMI e della comunità accademica, o potrebbe essere reso obsoleto dai nuovi sviluppi tecnologici. Si presume inoltre che a un regime normativo più severo corrispondano investimenti pubblici proprio per contrastare l'effetto di soffocamento dell'innovazione. Secondo la tradizione della previsione, si presume che il futuro sarà diverso da tutti e quattro gli scenari, ma potrà contenere elementi presenti in ciascuno di essi. Pertanto, considerati collettivamente con le loro storyline e la loro valutazione in termini di impatti potenziali su dimensioni chiave, gli scenari possono far emergere implicazioni e raccomandazioni rilevanti per i decisori politici.

TRAME DEGLI SCENARI

Innovazione guidata dal mercato (MDI).

In questo scenario la dinamica socioeconomica è inclusiva e il regime normativo leggero. L'assenza di una forte regolamentazione non provoca una posizione dominante da parte degli operatori

storici, perché la pluralità di attori in grado di rimanere sul mercato favorisce gli sviluppi tecnologici che abbassano le barriere in termini di dati e potenza di calcolo richiesta.

In assenza degli oneri amministrativi che accompagnano una regolamentazione pesante, le start-up, le scale-up, le PMI e la comunità della ricerca possono rimanere sul mercato accanto alle Big Tech e ai primi operatori già consolidati. I finanziamenti in capitale di rischio si diffondono in modo più capillare tra i diversi tipi di operatori e tra i vari Paesi, dando ulteriore impulso a un ecosistema di innovazione aperta guidato dalla logica di un mercato competitivo. La pluralità di operatori comporta anche un aumento e una diversificazione delle attività di R&S, con lo sviluppo costante di nuovi modelli senza il predominio e il blocco dei modelli introdotti sul mercato dai primi operatori.

In questo contesto, il numero di attori e talenti al lavoro contribuirà notevolmente al miglioramento della tecnologia stessa, e questa tendenza coinvolge sempre più entità europee, dalle grandi aziende, alle PMI, alle start-up e alla comunità accademica, il che è positivo per l'autonomia strategica. Il successo degli approcci open-source allo sviluppo e all'innovazione dei modelli potrebbe anche essere un fattore che contribuisce a rendere più rilevanti gli operatori europei. Quindi, la tecnologia stessa migliorerà, riducendo notevolmente i casi di "allucinazioni" ma anche il verificarsi effettivo del tipo di rischi sociali esaminati nella sezione 4.1, che contribuiscono alla coesione e all'equità sociale. Questo contributo sarà rafforzato dal fatto che le soluzioni della GAI, frutto di un'autoregolamentazione volontaria, saranno trasparenti e consentiranno di risolvere gli eventuali reclami dei consumatori. Inoltre, il potenziale successo nell'ulteriore sviluppo di approcci federati o di Small Language Models (SLM) e di tecniche di perfezionamento di modelli esistenti, come la tecnica Low-Rank Adaptation (LoRA), potrebbe ridurre i requisiti di dati e di potenza di calcolo rispetto agli approcci monolitici e contribuire a mantenere il nuovo mercato della GAI su un piano di parità. I miglioramenti tecnologici possono anche rendere più efficiente

il pre-addestramento e la messa a punto dei modelli utilizzando i dati disponibili al pubblico e quelli ottenuti dalle piattaforme esistenti, secondo le disposizioni del DMA. In questo scenario, i dati proprietari accumulati dagli operatori storici che gestiscono una piattaforma non rappresentano più un vantaggio insormontabile. Minori requisiti di dati e di potenza di calcolo riducono le opportunità e gli incentivi per pratiche di concorrenza sleale come il bundling (soprattutto con i servizi cloud) da parte delle Big Tech. Inoltre, anche se la regolamentazione è leggera, le autorità antitrust monitoreranno il comportamento concorrenziale degli operatori storici delle grandi tecnologie e impediranno l'adozione di tali pratiche.

Grazie al crescente numero di soluzioni introdotte sul mercato e alla presenza sempre più massiccia di numerosi operatori, non solo l'efficienza e l'aumento della produttività contribuiranno alla crescita complessiva, ma potenzialmente l'impatto sulla coesione sociale e sull'equità democratica sarà positivo. Con l'aumento delle applicazioni potenziali della GAI e la presenza di molti operatori sul mercato, i livelli di occupazione possono rimanere stabili o addirittura aumentare, sia nelle imprese che producono GAI sia in quelle che le utilizzano. In secondo luogo, il miglioramento dei modelli e della capacità di utilizzare meno dati ma di maggiore qualità può ridurre sia le "allucinazioni" sia la discriminazione sociale prodotta dall'algoritmo. I nuovi sviluppi tecnologici nel campo della GAI stessa e nella privacy differenziale possono ridurre i rischi per la privacy e per le violazioni dei dati. In terzo luogo, l'autoregolamentazione integrata da migliori soluzioni tecnologiche che blocchino il riutilizzo malevolo di contenuti, immagini e video generati dalla GAI ridurrebbe anche il rischio di deep fake e disinformazione che distorcono la democrazia e la sfera pubblica.

Questo è lo scenario ipotizzato in un contesto market-driven "puro". Tuttavia, occorre sottolineare come esista anche la possibilità che questo scenario consolidi il potere delle grandi aziende tecnologiche, piuttosto che promuovere una vera

concorrenza. Grandi attori come Google, Microsoft o Amazon non solo dominano il mercato, ma influenzano anche chi può entrare e come. Spesso, sono proprio queste aziende a finanziare le nuove startup o i nuovi operatori. Questo meccanismo di finanziamento, che sembra favorire la nascita di nuovi attori, in realtà genera un sistema dove i nuovi operatori restano strettamente legati ai finanziatori, limitando così la vera innovazione indipendente. Le startup nate sotto l'ala delle big tech diventano più delle aziende affiliate o consociate, piuttosto che dei veri concorrenti.

Innovazione guidata dal governo (GSI).

In questo scenario vi è la presenza di una regolamentazione rigorosa combinata con dinamiche socio-economiche inclusive, rese possibili da ingenti investimenti pubblici a sostegno di innovatori e start-up dell'AI. Uno scenario di questo tipo consente al governo di assumere un ruolo attivo nel plasmare il mercato, imponendo una regolamentazione e orchestrando investimenti concertati, ad esempio per stimolare le scale-up europee e fornire un'infrastruttura informatica pubblica che può produrre effetti positivi in termini di posti di lavoro e ricavi. Una forte regolamentazione può contribuire alla creazione di un mercato più inclusivo quando riduce le barriere all'ingresso di nuovi operatori, promuovendo la sperimentazione e l'esplorazione e consentendo agli operatori di crescere attraverso politiche di innovazione. Questo bilanciamento richiede una visione e una strategia da parte del governo per sviluppare una posizione tecnologica e industriale equilibrata e una politica di investimento mirata che integri le azioni normative, al fine di creare un mercato orchestrato. L'innovazione assume una forma di integrazione tra grandi operatori storici che incorporano in varie forme piccoli attori innovativi provenienti da PMI, start-up e dalla comunità accademica. Se in questo scenario, ad esempio, il governo europeo riesce a promuovere l'innovazione attraverso efficaci partenariati pubblico-privato, questo scenario consente di raggiungere un buon equilibrio tra regolamentazione e leadership industriale.

In questo scenario, la dinamica del mercato rimane inclusiva

senza chiusura monopolistica/oligopolistica per due motivi. In primo luogo, gli investimenti pubblici riescono a sostenere gli innovatori e le PMI più forti nel campo dell'AI, aiutandoli a competere e a rimanere sul mercato nonostante la presenza di forti operatori storici. In secondo luogo, le dinamiche di mercato invitano gli operatori storici ad aiutare quelli più piccoli ad affrontare la regolamentazione più rigida, beneficiando in cambio dei miglioramenti tecnologici e dell'accesso ai talenti. Come dimostra il caso della partnership di Mistral con Microsoft, per gli innovatori e le PMI europee collaborare con le Big Tech statunitensi è l'unico modo per crescere rapidamente. Quindi, rimane in qualche misura uno scenario inclusivo, perché le Big Tech non chiudono del tutto il mercato e permettono a molte iniziative ad esse collegate di svilupparsi e creare ulteriori posti di lavoro, e gli investimenti pubblici aiutano anche i piccoli operatori innovativi a rimanere sul mercato. Quindi, l'aumento dell'efficienza e della produttività non comporta una perdita di posti di lavoro. I grossi operatori beneficiano anche di miglioramenti tecnologici per ridurre le "allucinazioni" e l'effettivo verificarsi di rischi per la società. I governi possono anche orchestrare attivamente il mercato spezzando i monopoli delle Big Tech in entità più piccole, forzando una concorrenza equa sul mercato. In questo scenario le entità europee manterranno un loro spazio e avranno opportunità di crescita, ma in modo complementare rispetto allo scenario di innovazione aperta e coesiva.

Innovazione dominata dalle Big Tech (BDI).

Questo scenario ritrae una dinamica socioeconomica con un approccio normativo più pesante e severo che impone tutti gli obblighi in modo uniforme a ogni fornitore di modelli di fondazione, indipendentemente dalle dimensioni. Le Big Tech, grazie alla loro competenza in campo legale e alla loro capacità organizzativa, riusciranno a far fronte con successo a tale regime sia in termini di oneri amministrativi associati, sia trovando scappatoie e, quando necessario, impugnando le decisioni normative in tribunale. Al contrario, una regolamentazione pesante ridurrà seriamente la capacità di stare sul mercato di start-up, scale-up, PMI. Questo,

a sua volta, ridurrà la diversità degli attori e, di conseguenza, soffocherà l'innovazione tecnologica e lo sviluppo.

La riduzione del numero di attori coinvolti nell'introduzione di soluzioni e applicazioni nuove ed efficienti sul mercato significa che la tecnologia si svilupperà secondo linee più vantaggiose per i primi operatori e per le Big Tech, il che significa concentrarsi su modelli di fondazione che richiedono una quantità sempre maggiore di dati e di potenza di calcolo. Di conseguenza, il mercato sarà dominato da pochi operatori Big Tech che acquisiranno gli operatori iniziali di GAI relativamente grandi, e diventeranno i principali finanziatori/utilizzatori dell'innovazione prodotta dalla comunità di ricerca. Questo scenario può portare a posizioni direttamente contrapposte tra i player delle Big Tech e i governi, dove questi ultimi elaborano regolamenti sempre più severi per limitare le Big Tech. Paradossalmente, ciò può portare alla cattura della regolamentazione a favore delle Big Tech, lasciando ai governi pochi o nessuno strumento per orchestrare il mercato.

In questo scenario caratterizzato da una rigida regolamentazione e da un mercato dominato da poche Big Tech che operano in domini generici con modelli di fondazione monolitici, ci sono pochi incentivi per questi operatori storici a espandersi ulteriormente in nicchie o a spingere per lo sviluppo di modelli federati, o a continuare a investire nello sviluppo open-source, se questo non porta loro un vantaggio competitivo diretto rispetto agli altri grandi operatori, dato che la concorrenza dal basso verso l'alto da parte degli sfidanti sarà per lo più assente. Ciò può portare a una continuazione dell'attuale approccio monolitico, generico e closed-source alla costruzione di modelli di fondazione, nonostante l'esplorazione di approcci alternativi possa condurre a modelli complessivamente migliori.

A causa della mancanza di una pluralità di attori e talenti e per via degli interessi più restrittivi di pochi grandi operatori, la tecnologia migliorerà, ma non tanto quanto nello scenario dell'innovazione guidata dal mercato, soprattutto per quanto riguarda la riduzione

delle “allucinazioni” e il verificarsi del rischio sociale esaminato nella Sezione 4.1. Dato il loro potere lobbistico e legale, i grandi operatori tecnologici possono assumere un ruolo di opposizione e ricorrere ai tribunali quando i consumatori e/o i cittadini sollevano questioni di discriminazione sociale, disinformazione e privacy. Ciò non contribuirà positivamente alla coesione sociale e all'equità. A causa della concentrazione e del minor numero di soluzioni introdotte sul mercato, se da un lato la produttività cresce, dall'altro c'è il rischio di un calo dell'occupazione, perché questa ondata di innovazione crea molte più mansioni legate alla persona, indebolendo così la coesione sociale. Questo scenario sarà dominato dai giganti tecnologici statunitensi e dalle aziende cinesi, lasciando rimarrà poco spazio alle entità europee.

Innovazione oligopolistica (OI).

Questo scenario ritrae una dinamica socio-economica con un approccio normativo che consentirà a start-up, PMI e studiosi accademici di innovare la GAI senza dover rispettare obblighi e requisiti rigidi. In questo contesto inclusivo, il numero di operatori e talenti al lavoro contribuisce a migliorare la tecnologia stessa, in particolare in un contesto tecnologico in cui l'innovazione federata e open-source supera gli sforzi monolitici closed source. Questa tendenza potrebbe coinvolgere sempre più entità europee, dalle grandi aziende alle PMI, alle start-up e alla comunità accademica come nuovi attori del mercato. Tuttavia, per via della crescita precoce e del dominio del mercato, nonché dei metodi di concorrenza non adeguatamente controllati dall'antitrust, è probabile che i grandi operatori storici continuino a dominare il mercato, in quanto la scala di calcolo, di dati e di capitale a rischio rimane vantaggiosa. Questo può essere svantaggioso dal punto di vista europeo in termini di benefici economici, autonomia strategica, possibilità di dati e modelli di regolamentazione e supervisione dell'UE, in quanto affidare le infrastrutture crea dipendenze.

In un contesto in cui i grandi operatori storici sono in grado di rimanere competitivi e di attrarre risorse senza un approccio di open data e open science, tutta l'innovazione rilevante potrebbe avvenire dietro a fossati chiusi e i vantaggi degli operatori storici

si estenderanno ulteriormente. Gli operatori più piccoli, tra cui molte entità dell'UE, saranno quindi costretti a collaborare con questi grandi operatori storici. A causa della concentrazione e del minor numero di soluzioni introdotte sul mercato, se da un lato la produttività cresce, dall'altro c'è il rischio di un calo dell'occupazione, dato che questa ondata di innovazione automatizza molte più mansioni legate all'uomo.

È probabile che questo scenario rafforzi ulteriormente il dominio del mercato da parte dei giganti tecnologici statunitensi e delle aziende cinesi, con limitate opportunità competitive per le aziende e le iniziative europee.

VALUTAZIONE DEGLI SCENARI E CONCLUSIONI

La figura che segue presenta la valutazione dei quattro scenari rispetto a cinque dimensioni di impatto potenziale che spieghiamo di seguito.

In primo luogo, utilizziamo la crescita e ipotizziamo che essa derivi dai guadagni di produttività e di efficienza che l'innovazione della GAI produrrà. In secondo luogo, la coesione sociale si riferisce ai possibili effetti sui livelli di occupazione e sulla discriminazione sociale. In terzo luogo, nella voce equità democratica includiamo i possibili effetti della disinformazione (con impatto sul processo democratico) e della gestione scorretta della privacy. L'autonomia strategica misura la portata del ruolo di leader che le entità europee possono svolgere in ogni scenario, preservando l'autonomia del continente rispetto ai giganti tecnologici statunitensi e alle grandi aziende cinesi. Infine, la dimensione del contenimento del rischio valuta il grado di inserimento in ogni scenario di misure normative o di autoregolamentazione per ridurre/minimizzare i rischi descritti nella Sezione 4.1.

Riteniamo che la crescita sia leggermente superiore nel caso di Market Driven Innovation ("innovazione guidata dal mercato", d'ora in poi MDI) a causa del maggiore livello di concorrenza e di offerta. D'altra parte, presumiamo un contributo alla crescita elevato anche per lo scenario Big Tech Dominated Innovation ("innovazione guidata dalle Big Tech", d'ora in poi BDI) e gli altri due scenari Government Steered Innovation ("innovazione guidata dal governo", d'ora in poi GSI) e Oligopolistic Innovation ("innovazione oligopolistica", d'ora in poi OI). Presumiamo che questa tecnologia, in una forma o nell'altra, avrà un impatto positivo sulla crescita in tutti gli scenari, ma che gli scenari differiranno in relazione

alle altre dimensioni. La MDI è chiaramente superiore alla BDI in termini di coesione sociale, equità e autonomia strategica, mentre si comporta peggio in termini di gestione legalmente conforme del rischio a causa dell'incertezza che comporta un approccio più flessibile alla regolamentazione. La Government Steered Innovation, d'ora in poi GSI, ha ottenuto un punteggio leggermente migliore in termini di coesione sociale, equità, e autonomia strategica rispetto alla BDI, mentre quest'ultimo e l'OI ottengono punteggi molto simili.

Dalla discussione presentata nel report si traggono le seguenti conclusioni di rilevanza politica sia a livello di UE che di Italia.

CONCLUSIONI DELL'UE

Il futuro dell'AI generativa in Europa presenta uno scenario caratterizzato da un potenziale di miglioramento, sfide normative e dalla ricerca di una leadership di settore. Nel momento in cui l'Europa naviga nelle dinamiche evolutive dell'AI generativa, diversi fattori chiave possono essere ricavati dall'analisi e dagli scenari presentati in questo report:

- L'AI generativa promette di aumentare la produttività e di migliorare la qualità degli output e, potenzialmente, quella del lavoro in diverse mansioni. Sebbene permangano incertezze sulla dislocazione dei posti di lavoro e sulle implicazioni sociali, le prime indicazioni suggeriscono impatti positivi sulla crescita economica e sulle funzioni lavorative
- Attualmente l'Europa è in ritardo, rispetto alla concorrenza mondiale, nel campo dell'AI generativa con una significativa disparità nell'assegnazione di finanziamenti alle start-up europee rispetto alle controparti statunitensi e cinesi.

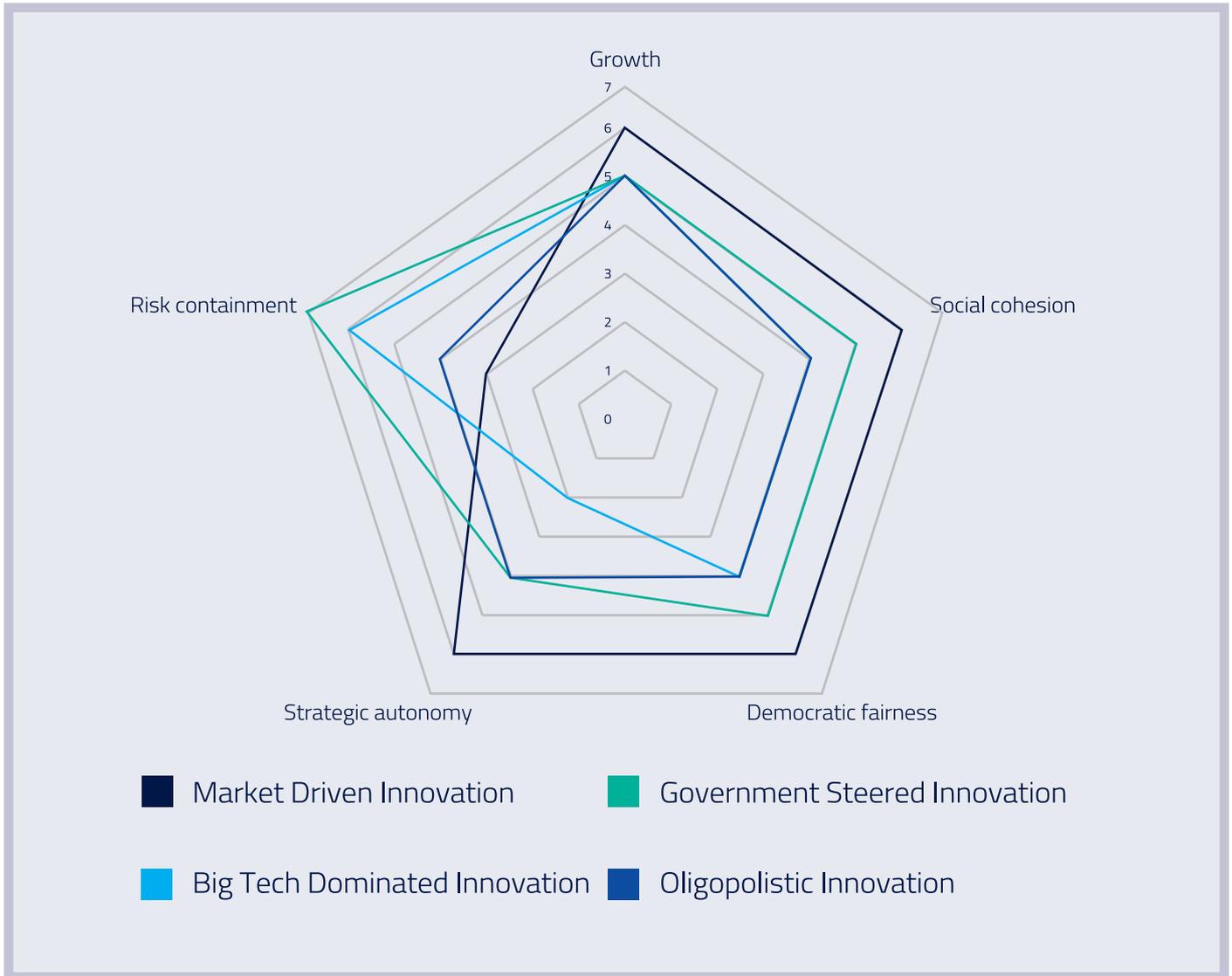


Figura 13 Valutazione a diagramma radar: dimensioni oggettive Fonte: Elaborazione propria

L'assenza di un'iniziativa unitaria riguardo l'AI a livello europeo pone delle sfide all'Europa per affermarsi con prodotti e infrastrutture di AI "Made in Europe".

- Oltre alle implicazioni economiche più tangibili, senza una posizione forte nel campo della GAI l'autonomia strategica dell'Europa si indebolirebbe e, allo stesso tempo, i Paesi europei dovrebbero accettare e adattarsi a soluzioni GAI culturalmente ed eticamente radicate in altre parti del mondo. Questo, tra l'altro, potrebbe avere un impatto sulla ricchezza linguistica dell'Europa, perché anche quando le soluzioni GAI saranno adattate avranno come riferimento l'inglese e perderanno le sfumature linguistiche delle molte altre lingue europee.
- L'aumento dei finanziamenti è essenziale, ma dovrebbe essere mirato a ecosistemi aperti e alla collaborazione, non solo a singoli grandi player tecnologici. Iniziative come quella proposta di GenAI4EU, per creare centri intersettoriali, possono contribuire a rendere più democratico l'accesso alle risorse. I progetti rivelati dalla Commissione Europea, il 24 gennaio 2024, per supportare le start-up e le PMI di Intelligenza Artificiale sono un primo passo, ma c'è ancora molto da fare e i finanziamenti concessi non sono sufficienti.
- Non si tratta peraltro soltanto di incentivare i finanziamenti pubblici. In Italia e in Europa, l'aspettativa è spesso che sia il governo, attraverso investimenti pubblici, a guidare la creazione di innovazioni tecnologiche, come l'intelligenza artificiale generativa. Anche negli Stati Uniti l'apparato governativo spesso sostiene le aziende tecnologiche, ad esempio tramite contratti miliardari assegnati dal Dipartimento della Difesa ad aziende che usano la AI per la sorveglianza, tuttavia l'esperienza statunitense dimostra che molte grandi aziende tech sono state in grado di crescere grazie a ingenti investimenti privati, senza fare affidamento sul sostegno pubblico. Questo modello ha permesso agli Stati Uniti di emergere come leader mondiali nel campo dell'AI. L'Europa, e in particolare l'Italia, dovrebbero dunque adottare un approccio simile, creando un ambiente che favorisca gli

investimenti privati nelle grandi imprese, riducendo al minimo le barriere burocratiche e facilitando l'accesso al capitale.

- Basarsi solo sulla regolamentazione e sul previsto effetto Bruxelles che ne deriva non è chiaramente sufficiente a colmare le attuali lacune. La regolamentazione è importante per creare certezza e garantire la riduzione dei rischi per la società, ma senza una strategia di sostegno agli investimenti potrebbe potenzialmente aumentare ulteriormente le lacune esistenti. I responsabili politici sono invitati ad adottare misure proattive per trovare un punto di equilibrio tra l'innovazione e le considerazioni etiche. Le piattaforme di collaborazione possono facilitare il dialogo sulla mitigazione dei rischi e la massimizzazione delle opportunità associate alle tecnologie dell'AI. Se da un lato la regolamentazione è fondamentale per mitigare i rischi, dall'altro un suo eccesso può soffocare l'innovazione, soprattutto per le start-up più piccole. Sarebbe utile realizzare sandbox e collaudi regolamentari per i ricercatori e le aziende europee di AI generativa al fine di sperimentare in modo responsabile in conformità alla legge europea sull'AI.
- Oltre alle iniziative volte alla regolamentazione come l'AI Act, è necessario affrontare le incertezze relative alle politiche di concorrenza e alle iniziative antitrust nel quadro normativo dell'AI generativa. Il rischio di pratiche di concorrenza sleale da parte degli operatori storici nel mercato dell'AI generativa, come l'accorpamento di prodotti o la stipula di accordi in esclusiva, deve essere mitigato. Le potenziali barriere all'accesso per i nuovi operatori devono essere abbassate favorendo l'accesso al talento e alla potenza di calcolo.
- Come evidenziato nel report "Il futuro della competitività europea", presentato alla Commissione Europea lo scorso 9 settembre dall'ex premier italiano ed ex presidente della Banca Centrale Europea Mario Draghi, l'Europa ha accumulato un ritardo significativo in alcune aree chiave dell'innovazione tecnologica, e colmare questo divario richiederà tempo e strategie mirate. Il focus dovrebbe essere su come navigare in un contesto che ha vincoli giuridici e strutturali unici,

sfruttando le risorse e le opportunità che già esistono. Non è realistico aspettarsi che l'Europa possa produrre una "Google europea" in tempi brevi, ma è possibile individuare spazi strategici dove l'innovazione possa prosperare, lavorando su iniziative che massimizzino il potenziale delle piccole e medie imprese, delle startup e della ricerca scientifica.

Approfondendo gli aspetti più specifici, presentiamo le seguenti cinque conclusioni:

1. **La regolamentazione come co-creazione.** Il lavoro sulla regolamentazione della GAI, compresa la messa in atto della legislazione attraverso standard, norme e vigilanza normativa, dovrebbe continuare, ma coinvolgendo meglio le industrie emergenti in un processo di co-creazione basato sull'apprendimento attraverso la pratica e sulla costante ottimizzazione attraverso sandbox e altre forme di sperimentazione.
2. **Approccio orientato alla missione a livello europeo.** Il futuro della GAI ha un impatto sulla sovranità europea e pertanto deve essere parte integrante di una strategia globale dell'UE, piuttosto che affidarsi a iniziative scollegate degli Stati membri. È necessario istituire una politica industriale coordinata della GAI con meccanismi di finanziamento ad hoc per sostenere e collegare i poli di eccellenza già esistenti in Europa. L'attenzione dovrebbe essere olistica, guidata da una visione a lungo termine sull'AI, portare all'eccellenza competitiva a livello globale e comprendere lo sviluppo di modelli, l'infrastruttura informatica, i dati, le competenze, la ricerca e l'innovazione.
3. **Nuovi strumenti di R&S.** Gli strumenti tradizionali di R&S non funzionano nel mondo in rapida evoluzione dell'AI; questi andrebbero modernizzati e maggiormente orientati ai risultati. Una vera e propria sfida di AI generativa su larga scala, ad esempio, dovrebbe essere organizzata con un budget di almeno 0,5 miliardi di euro.
4. **High Performance Computing.** L'impresa comune europea per il calcolo ad alte prestazioni (European High Performance Computing Joint Undertaking, EuroHPC JU) ha avuto finora un

impatto molto positivo in Europa. L'attenzione, tuttavia, si è concentrata soprattutto sull'HPC per le cosiddette "scienze dure" (fisica, chimica, biologia, medicina, ecc.), mentre è necessaria un'iniziativa HPC rivolta specificamente agli informatici, agli sviluppatori di modelli di GAI e all'industria.

5. **Oltre gli LLM. La GAI non si limita a gestire gli LLM.** Occorre compiere maggiori sforzi per coniugare i modelli di fondazione con tecniche di modellazione semantica per supportare una modellazione più esplicita dello stato di ragionamento, ad esempio per consentire spiegazioni del ragionamento, nonché sforzi e ricerche su architetture alternative di modelli di fondazione, come modelli federati specializzati su scala ridotta.

CONCLUSIONI/RACCOMANDAZIONI PER IL CONTESTO ITALIANO

Molte delle conclusioni e delle raccomandazioni di cui sopra, considerate da una prospettiva europea, sono applicabili anche alla situazione dell'Italia. Tuttavia, aggiungiamo anche alcune raccomandazioni specifiche per il contesto italiano.

Il potenziale impatto dell'AI Generativa sulla produttività del Sistema-Italia è straordinario. Secondo un modello di stima elaborato da The European House - Ambrosetti, l'implementazione di queste soluzioni potrebbe incrementare il PIL italiano fino al 18%, rappresentando un'opportunità senza precedenti per rafforzare l'economia nazionale. In termini concreti, questo potrebbe tradursi in un aumento del valore aggiunto annuo di 312 miliardi di euro o in una riduzione delle ore di lavoro di 5,7 miliardi, a parità di valore aggiunto prodotto. Un valore paragonabile a 1,6 volte l'intero Piano Nazionale di Ripresa e Resilienza (PNRR) o al PIL della Lombardia. Tuttavia, l'adozione dell'AI Generativa non è priva di rischi etico-sociali che devono essere affrontati con un approccio responsabile. La trasparenza nell'uso dei dati, l'affidabilità delle tecnologie implementate, la sicurezza e la privacy dei sistemi, l'equità negli output generati e l'inclusività nell'accesso a questi servizi sono aspetti fondamentali per uno sviluppo etico dell'AI. È essenziale

mantenere un coinvolgimento umano nel processo decisionale e nella creazione di contenuti derivanti da strumenti di AI Generativa, in un'ottica di "human-in-the-loop". Questo richiede un dialogo costruttivo tra governi, industria

Attualmente, l'Italia, nel contesto europeo dell'intelligenza artificiale presenta una posizione complessa. Nonostante l'innovazione e la ricerca nel paese stiano crescendo - secondo una ricerca del Politecnico di Milano, il mercato dell'AI in Italia è cresciuto nel 2023 del 52%, raggiungendo i 760 milioni di euro - esistono significative disparità rispetto ai leader del settore come la Francia, noti per i loro avanzati investimenti strategici, un dinamico ecosistema di startup e un forte sostegno governativo nel campo dell'AI. Per ridurre questo divario e aspirare a un livello di eccellenza comparabile a quello francese, l'Italia deve adottare un approccio più strutturato e integrato.

Il governo italiano ha presentato ad aprile 2024 un disegno di legge in materia di AI, un passo importante verso la regolamentazione di questo settore in rapida crescita. Tuttavia, come spesso accade, il percorso legislativo è lungo e incerto, con il disegno di legge ancora all'esame delle camere. Per quanto riguarda il finanziamento pubblico, il governo ha proposto di stanziare un miliardo di euro, una cifra che, sebbene significativa, viene considerata insufficiente per affrontare le sfide su larga scala che l'AI presenta, soprattutto in un mercato dominato da giganti come gli Stati Uniti e la Cina. Inoltre, l'attenzione si concentra su piccole e medie imprese (PMI) e startup, che costituiscono la spina dorsale dell'economia italiana. Tuttavia, le esigenze finanziarie di tali realtà, spesso più fragili e ad alto rischio, richiedono un supporto più consistente per affrontare il costo della ricerca e dello sviluppo.

Il primo passo essenziale per l'Italia consiste quindi nell'aumentare gli investimenti pubblici e privati specificatamente destinati alla ricerca e allo sviluppo in AI. È fondamentale non solo incrementare il finanziamento, ma anche orientarlo verso progetti che possano concretamente stimolare l'interesse sia industriale che accademico, creando un ciclo virtuoso di innovazione e applicazione

pratica. Parallelamente, la creazione di centri di eccellenza e hub di innovazione potrebbe servire come catalizzatore per il trasferimento tecnologico dall'accademia all'industria. Questi centri dovrebbero essere strategicamente collegati alle università e agli istituti di ricerca, fornendo al contempo solide connessioni con le industrie italiane ed europee, favorendo così un ambiente in cui l'innovazione può prosperare.

Per supportare ulteriormente l'innovazione, l'Italia dovrebbe implementare politiche incentrate sulle startup e sulle piccole e medie imprese che operano nel campo dell'AI. Attraverso agevolazioni fiscali, accesso facilitato a finanziamenti e programmi di accelerazione, queste imprese possono crescere e competere a un livello internazionale. La creazione di un ambiente imprenditoriale favorevole è cruciale per stimolare un ecosistema di GAI competitivo e innovativo.

Un tema chiave è quello della fiducia. Il progresso tecnologico richiede investimenti e un ambiente che permetta la crescita delle idee, anche quando non si conosce in anticipo il risultato finale. Guardando al passato, un esempio emblematico è quello di Guglielmo Marconi, che scoprì la radio senza sapere in anticipo dove l'avrebbero portato le sue ricerche. Marconi non aveva garanzie, ma disponeva di un contesto che gli consentì di esplorare e innovare. L'Italia (e l'Europa) deve investire nella creazione di un ecosistema che favorisca la fiducia, sostenga le aziende nella loro crescita e attragga capitali. Senza questi elementi fondamentali, sarà difficile competere con ecosistemi come quello statunitense, dove la fiducia nel potenziale innovativo e l'integrazione tra governo, aziende e ricerca accademica creano le condizioni ideali per lo sviluppo di tecnologie di punta.

Occorre inoltre semplificare i processi burocratici e ridurre gli ostacoli normativi. Molti imprenditori, fra quali alcuni sentiti in fase di revisione di questo report, riscontrano che in Italia esistono barriere che non trovano in altre nazioni europee. Burocrazia, incertezze normative e regolamentazioni poco chiare rendono difficile

l'operatività delle imprese. Queste barriere devono essere eliminate per incoraggiare le aziende italiane a investire in innovazione e competere con i grandi player internazionali.

L'educazione e la formazione rappresentano un altro pilastro fondamentale.

L'automazione e l'intelligenza artificiale stanno trasformando rapidamente il mercato del lavoro, e molti posti di lavoro tradizionali sono destinati a scomparire. Invece di concentrare gli sforzi nella protezione di questi lavori, l'Italia e l'Europa dovrebbero investire nella formazione di nuove competenze. È necessario preparare i lavoratori per i nuovi ruoli che emergeranno grazie alle innovazioni tecnologiche, seguendo l'esempio statunitense, dove la riqualificazione e la flessibilità del mercato del lavoro sono più sviluppate.

Investire nella formazione e nello sviluppo delle competenze digitali diventa cruciale per preparare la forza lavoro all'utilizzo di soluzioni di AI Generativa. Si stima che all'Italia mancherebbero 3,7 milioni di occupati con competenze digitali di base e 137.000 iscritti in più a corsi di laurea ICT per abilitare l'implementazione di queste tecnologie nel tessuto economico italiano. Ampliare l'offerta formativa nelle università e negli istituti tecnici è essenziale per garantire che i futuri professionisti acquisiscano le competenze richieste nel settore. Programmi di formazione continua possono inoltre aiutare i professionisti attuali a rimanere aggiornati sulle ultime tecnologie e metodologie.

Una volta formate le persone, occorre però trovare il modo di trattenerle sul territorio – quello della “fuga dei cervelli” è un problema reale - e specularmente di attrarre talenti dall'estero con iniziative mirate.

Un modello a cui ispirarsi potrebbe essere quello degli Usa, che hanno introdotto un visto veloce per i professionisti dell'AI, facilitando l'immigrazione di esperti qualificati negli Stati Uniti. In risposta a queste dinamiche l'Europa, e l'Italia in particolare,

dovrebbero rendere più agevole l'arrivo di talenti dall'estero, riducendo le barriere burocratiche e offrendo incentivi competitivi per trattenere i talenti.

In conclusione, l'Intelligenza Artificiale Generativa rappresenta un'opportunità straordinaria per l'Italia di superare le sfide dell'invecchiamento della popolazione e della stagnazione della produttività, ma richiede un approccio olistico che affronti le questioni etiche, la digitalizzazione delle imprese e lo sviluppo delle competenze digitali. Solo attraverso un impegno coordinato tra governi, aziende e società civile, l'Italia potrà cogliere appieno i benefici di questa rivoluzione tecnologica, mantenendo al contempo un legame umano essenziale nell'evoluzione del mondo del lavoro e preservando i valori fondamentali di equità, inclusività e responsabilità.

RINGRAZIAMENTI

Questo studio, intitolato “AI Generativa: la ricerca di una regolamentazione e di una leadership industriale in Europa”, fa parte della serie di report Makers & Shapers di EIT Digital. Questi rapporti affrontano aspetti specifici delle tecnologie e degli sviluppi digitali. Seguono un approccio basato su scenari, fondato sugli sviluppi e sullo stato dell’arte nel settore specifico, e analizzano l’impatto dei diversi scenari su una serie di indicatori predefiniti. Questo report è stato sviluppato in collaborazione con AiNed e con il supporto della EIT AI Community coordinata da EIT Digital.

Cristiano Codagnone è stato incaricato di supportare lo studio e ha scritto il rapporto sotto la guida e la supervisione di EIT Digital e AiNed. Coadiuvato da Giovanni Liva, Codagnone ha fornito ampiezza e profondità allo studio grazie a discussioni interdisciplinari con gli stakeholder e a un’ampia analisi e indagine della letteratura. Lo ringraziamo per il suo contributo. Un ringraziamento speciale va a Stefan Leijnen di AiNed. Infine, vorremmo ringraziare i membri della nostra commissione di esperti che si sono riuniti il 6 e 7 marzo 2024 e l’11 settembre 2024 per aver revisionato il rapporto e aver fornito un contributo prezioso a questo studio:

Daniel Abbou, Managing Director, German AI Association; Slavina Ancheva, Team Leader and Policy Adviser to Member of the European Parliament Eva Maydell, European Parliament; Brando Benifei, Member of the European Parliament; Francesco Bonfiglio, CEO, Dynamo; Cristina Brandtstetter, Chief Marketing Officer, Dynamo; Igor Carron, CEO and Co-founder, Lighton; Professor Dominic Chalmers, University of Glasgow; Edwin M. Colella, Chief Sales and Marketing Officer, Wallife; Sabine Demey, Program Director, Flanders Artificial Intelligence Research Program; Marta Fasan, Ricercatrice, Dipartimento di Giurisprudenza dell’Università di Trento; Emanuela Girardi, President, Adra - AI, Data and Robotics association; Frauke Goll, Managing Director, Applied AI; Jeannette

Gorzala, Vice President, European AI Forum; Nicola Grandis, CEO, ASC27; Mirko Guarnera, AI and Personal Electronics Marketing manager, STMICROELECTRONICS; Arash Hajikhani, Research Team Leader – Quantitative science and technology studies, VTT - Technical Research Centre of Finland; Professor Fredrik Heintz, Linköping University; Professor of Artificial Intelligence and Computer Science Tom Heskes, Radboud University; Professor of Artificial Intelligence and Computer Science Holger Hoos, RWTH Aachen University; Philipp Hutchinson, Senior AI Strategist, Applied AI; Vangelis Karkaletsis, Director of the Institute of Informatics & Telecommunications, NCSR Demokritos; Benjamin Larsen, Lead - Artificial Intelligence and Machine Learning, World Economic Forum; Lorenzo Maternini, Membro del board, Comitato governativo sull’Intelligenza Artificiale; George Metakides, President, Digital Enlightenment Forum; Francesco Musso, CEO, Dedalo.ai; Cosimo Musca, Deputy Head of Italy Public Affairs, STMICROELECTRONICS; Tero Ojanperä, Co-Founder and Executive Chair, Silo AI; Riccardo Osti, CEO, Wonderflow.ai; Bálint Árpád Pataki, Accredited Parliamentary Assistant to the Member of the European Parliament Eva Maydell, European Parliament; Jim Spohrer, Member of the Board of Directors of the non-profit ISSIP and ServCollab; Niels Taatgen, Research Director, University of Groningen; Orestis Trasanidis, Coordinator, EIT AI Community; Kai Zenner, Head of Office and Digital Policy Adviser to Member of the European Parliament Axel Voss, European Parliament; Onno Zoeter, Principal Data Scientist, Booking.com.

Bruxelles, marzo 2024 – Trento, ottobre 2024

BIBLIOGRAFIA/SITOGRAFIA

- Knott, A., Pedreschi, D., Chatila, R. et al. (2023). Generative AI models should include detection mechanisms as a condition for public release. *Ethics Inf Technol* 25, 55 (<https://doi.org/10.1007/s10676-023-09728-4>).
- Eloundou, T., S. Manning, P. Mishkin and D. Rock (2023) 'GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models', mimeo, available at <http://arxiv.org/abs/2303.10130>
- Hatzius, J., J. Briggs, D. Kodnani e G. Pierdomenico (2023) The Potentially Large Effects of Artificial Intelligence on Economic Growth, *Global Economics Analyst*, 26 marzo, Goldman Sachs Economics Research
- Ambrosetti & Microsoft. (2023). AI 4 Italy: impatti e prospettive dell'intelligenza artificiale generativa per l'Italia e il Made in Italy (<https://www.ambrosetti.eu/news/ai-4-italy-impatti-e-prospettive-dellintelligenza-artificiale-generativa-per-litalia-e-il-made-in-italy/>).
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1-38.
- OCSE (2023). AI Language Models: Technological, Socio-Economic and Policy Consideration', *OECD Digital Economy Papers* 352 (<https://doi.org/10.1787/13d38f92-en>).
- In sei esperimenti, i partecipanti (N = 4.600) non sono stati in grado di rilevare le autopresentazioni generate da modelli linguistici AI all'avanguardia in contesti professionali, di ospitalità e di appuntamenti, si veda Jakesch, M., Hancock, J. T., & Naaman, M. (2023). Human heuristics for AI-generated language are flawed. *Proceedings of the National Academy of Sciences*, 120(11), e2208839120. Risultati simili sono riportati anche in Waltzer, T., Cox, R. L., & Heyman, G. D. (2023). Testing the Ability of Teachers and Students to Differentiate between Essays Generated by ChatGPT and High School Students. *Human Behavior and Emerging Technologies*, 1923981, doi:10.1155/2023/1923981
- Knott, A., Pedreschi, D., Chatila, R. et al. (2023). Generative AI models should include detection mechanisms as a condition for public release, op. cit., pag. 2.
- Wang, C., Liu, S., Yang, H., Guo, J., Wu, Y., & Liu, J. (2023). Ethical Considerations of Using ChatGPT in Health Care. *J Med Internet Res*, 25, e48009. doi:10.2196/48009.
- Nel marzo 2023, il primo resoconto del Parlamento anticipava ciò che sarebbe avvenuto in seguito: cfr. Madiega, T. General Purpose Artificial Intelligence, EPRS | European Parliamentary Research Service, marzo 2023 ([https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/745708/EPRS_ATA\(2023\)745708_IT.pdf](https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/745708/EPRS_ATA(2023)745708_IT.pdf)). Il documento sottolinea i rischi dell'AI generativa e la necessità di regolamentarla. L'intero processo legislativo dell'AI Act e la nuova prospettiva sono stati poi presentati in un'altra sintesi pubblicata nel giugno 2023: cfr. Madiega, T., EU Legislation in Progress: Artificial Intelligent Act, EPRS | European Parliamentary Research Service, giugno 2023 ([https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_IT.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_IT.pdf)).
- Si veda: Madiega, T., Parliament's negotiating position on the artificial intelligence act, EPRS | European Parliamentary Research
- <https://arxiv.org/abs/1706.03762>
- <https://www.asimovinstitute.org/overview-neural-network-zoo/>
- Rae, J.W., S. Borgeaud, T. Cai, K. Millican, J. Hoffmann, F. Song ... G. Irving (2021). Scaling Language Models: Methods, Analysis & Insights from Training Gopher (<http://arxiv.org/abs/2112.11446>).
- Schick, T. and H. Schütze,(2021). 'It's Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners', *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (<https://doi.org/10.18653/v1/2021.naacl-main.185>).
- Hu, E.J., Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang and W. Chen (2021). LoRA: Low-Rank Adaptation of Large Language Models (<http://arxiv.org/abs/2106.09685>). Vale la pena di menzionare anche QLoRA, una tecnica più recente per una messa a punto efficiente. Riduce in modo significativo i requisiti di memoria per la messa a punto, quantizzando i pesi del modello in bit di dimensioni inferiori.
- Gli sviluppatori hanno utilizzato il modello open-source LLaMA e lo hanno messo a punto su dataset open-source di alta qualità contenenti dati di dialogo provenienti dal web, compreso il dialogo con i LLM, come ChatGPT. Si vedano: Geng, X., A. Gu-diband, H. Liu, E. Wallace, P. Abbeel, S. Levine and D. Song (2023). Koala: A Dialogue Model for Academic Research', *Berkeley Artificial Intelligence Research*, 3 April (<https://bair.berkeley.edu/blog/2023/04/03/koala/>).
- Dati riportati in Carugati, C. (2023). Competition in generative artificial intelligence

- foundation models, Working Paper 14/2023, Bruegel (<https://www.bruegel.org/sites/default/files/2023-09/WP%2014.pdf>), p. 7.
19. Carugati, C. (2023). Competition in generative artificial intelligence foundation models, op. cit., pag. 7.
 20. Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., Wei, J., Chung, H. W., ... & Natarajan, V. (2022). Large language models encode clinical knowledge. arXiv preprint arXiv:2212.13138.
 21. Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., ... & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper, (24-013).
 22. Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Disponibile su SSRN 4375283.
 23. Peng, S., Kalliamvakou, E., Cihon, P., & Demirel, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. arXiv preprint arXiv:2302.06590.
 24. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation, op.cit.
 25. Mons, B. (2023). Does the hype of Generative AI need top-down regulation, or will it implode? Frontiers | Policy Labs (<https://policylabs.frontiersin.org/content/commentary-does-the-hype-of-generative-ai-need-top-down-regulation-or-will-it-implode>).
 26. Ibidem.
 27. Ibidem.
 28. <https://a16z.com/who-owns-the-generative-ai-platform/>
 29. Fonte: Dealroom.co.
 30. <https://www.economist.com/the-world-ahead/2023/11/13/generative-ai-will-go-mainstream-in-2024>
 31. <https://info.kpmg.us/news-perspectives/technology-innovation/kpmg-usexecutives-genai-2023.html>
 32. <https://iot-analytics.com/leading-generative-ai-companies/>
 33. <https://www.reuters.com/technology/microsofts-deal-with-mistral-ai-faces-eu-scrutiny-2024-02-27/>.
 34. Cfr. comunicato stampa della CMA del 4 maggio 2023 (<https://www.gov.uk/government/news/cma-launches-initial-review-of-artificial-intelligence-models>).
 35. Si veda Sheppard Mullin Richter & Hampton LLP, AI Under the Antitrust Microscope: Competition Enforcers Focusing on Generative AI from All Angles, August 9 2023, Lexology (<https://www.lexology.com/library/detail.aspx?g=5604adda-62c0-4ffe-8007-94b802d4b81c>).
 36. Dickerson, J. e Novak, A. La presidente della FTC, Lina Khan, dice che l'AI potrebbe «mettere il turbo» alle frodi, essere usata per «schiacciare la concorrenza», CBS News 27 luglio 2023 (<https://www.cbsnews.com/news/lina-khan-ftc-chair-federal-trade-commission-ai-turbocharge-fraud-squash-competition/>).
 37. Federal Trade Commission (FTC), Generative AI Raises Competition Concerns, giugno 29, 2023 (<https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns>).
 38. Federal Trade Commission (FTC), Generative AI Raises Competition Concerns, op. cit.
 39. Carugati, C. (2023). Competition in generative artificial intelligence foundation models, op. cit.
 40. Come riportato in Carugati (2023), circa 37.000 dataset open-source sono disponibili su Hugging Face (<https://huggingface.co/datasets?sort=downloads>)
 41. Chui, M. et al (2023). The economic potential of generative AI. McKinsey & Company (<https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/the%20economic%20potential%20of%20generative%20ai%20the%20next%20productivity%20frontier/the-economic-potential-of-generative-ai-the-next-productivity-frontier-vf.pdf?shouldIndex=false>).
 42. Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. arXiv preprint arXiv:2303.10130.
 43. Peng, S., Kalliamvakou, E., Cihon, P., & Demirel, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. arXiv preprint arXiv:2302.06590.
 44. Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., ... & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper, (24-013).
 45. Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work (No. w31161). National Bureau of Economic Research.
 46. Eisfeldt, A. L., Schubert, G., & Zhang, M. B. (2023). Generative ai and firm values (No. w31222). National Bureau of Economic Research.
 47. McKinsey & Company. (2023). The state of AI in 2023: Generative AI's breakout year (<https://www.mckinsey.com/~media/mckinsey/business%20functions/quantumblack/our%20insights/the%20state%20of%20ai%20in%202023%20gener>

- ative%20ais%20breakout%20year/the-state-of-ai-in-2023-generative-ais-breakout-year-v3.pdf?shouldIndex=false).
48. <https://www.mckinsey.com/industries/healthcare/our-insights/tackling-health-cares-biggest-burdens-with-generative-ai>
 49. Wang, C., Liu, S., Yang, H., Guo, J., Wu, Y., & Liu, J. (2023). Ethical Considerations of Using ChatGPT in Health Care, op. cit.
 50. <https://www.latitudemedia.com/news/seven-ways-utilities-are-exploring-ai-for-the-grid>
 51. Bender, E. M., Gebru, T., McMillan-Major, A. e Shmitchell, S. (2021, marzo). On the dangers of stochastic parrots: Can language models be too big?. In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency (pp. 610-623).
 52. <https://www.theverge.com/2020/12/5/22155985/paper-timnit-gebru-fired-google-large-language-models-search-ai>
 53. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... e Liang, P. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.
 54. Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P. S., ... e Gabriel, I. (2021). Ethical and social risks of harm from language models. arXiv preprint arXiv:2112.04359.
 55. OCSE (2023). AI language models. Technological, socio-economic and policy considerations. OECD Digital Economy Papers.
 56. Lucy, L. e Bamman, D. (2021). Gender and representation bias in GPT-3 generated stories. In Proceedings of the Third Workshop on Narrative Understanding.
 57. Università di Stanford (2023): https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf
 58. Kremer, A. et al (2023). As gen AI advances, regulators—and risk functions—rush to keep pace. McKinsey & Company (https://www.mckinsey.com/~media/mckinsey/business%20functions/risk/our%20insights/as%20gen%20ai%20advances%20regulators%20and%20risk%20functions%20rush%20to%20keep%20pace/as-gen-ai-advances-regulators-and-risk-functions-rush-to-keep-pace_final.pdf?shouldIndex=false).
 59. Drake, M. et al. From Washington to Brussels: A Comparative Look at the Biden Administration's Executive Order and the EU's AI Act, October 30 2023, Inside Privacy (<https://www.insideprivacy.com/artificial-intelligence/from-washington-to-brussels-a-comparative-look-at-the-biden-administrations-executive-order-and-the-eus-ai-act/>).
 60. <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>
 61. Cfr. <https://www.techpolicy.press/the-many-questions-about-indias-new-ai-advisory/> e <https://dig.watch/updates/indias-it-ministry-issues-advisory-on-approval-and-labelling-of-ai-tools>
 62. <https://www.un.org/en/ai-advisory-body>
 63. <https://gpai.ai/>
 64. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5379
 65. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX-:52021PC0206>.
 66. <https://ainowinstitute.org/publication/gpai-is-high-risk-should-not-be-excluded-from-eu-ai-act>
 67. <https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf>
 68. Ibid., Art 3(1b).
 69. https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf.
 70. Ibid., Article 3(1c).
 71. <https://oecd.ai/en/wonk/foundation-models-eu-ai-act-fairer-competition>
 72. Quanto segue si basa sul contenuto del comunicato stampa del Parlamento europeo (<https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>), poiché il testo dell'accordo non è ancora stato pubblicato.
 73. Cfr.: <https://artificialintelligenceact.com/updates/#:-:text=Representatives%20from%20member%20states%20unanimously,agreement%20reached%20in%20December%202023>.
 74. <https://digital-strategy.ec.europa.eu/en/policies/ai-pact>
 75. Secondo Kai Zenner nella sua conferenza alla citata Digital Humanism Lecture Series del 19 dicembre 2023. (<https://www.youtube.com/digitalhumanism>).
 76. George Metakides nella sua conferenza alla citata Digital Humanism Lecture Series del 19 dicembre 2023. (<https://www.youtube.com/digitalhumanism>).
 77. Questa è parte di una dichiarazione richiesta da George Metakides e letta nella sua conferenza alla citata Digital Humanism Lecture Series (<https://www.youtube.com/digitalhumanism>). Anu Bradford ha espresso opinioni simili anche in varie altre dichiarazioni pubbliche, come ad esempio "EU Establishes World-Leading AI Rules, Could That Affect Everyone?"; VOA, 11 dicembre 2023 (<https://www.voanews.com/a/eu-establishes-world-leading-ai-rules-could-that-affect-every->

- one-/7393780.html).
78. https://www.bruegel.org/analysis/adapting-european-union-ai-act-deal-generative-artificial-intelligence#footnote4_30gc309
 79. Si vedano gli esempi di Stanford e dell'OCSE disponibili qui: <https://crfm.stanford.edu/2023/06/15/eu-ai-act.html#fn:1> e qui <https://oecd.ai/en/wonk/foundation-models-eu-ai-act-fairer-competition>.
 80. Soete, L. (2023). FAIR use in Artificial Intelligence? Access to data for the benefit of all, *Frontiers | Policy Labs* (<https://policylabs.frontiersin.org/content/commentary-fair-use-in-artificial-intelligence-access-to-data-for-the-benefit-of-all>).
 81. Ibid.,
 82. Mons, B. (2023). Does the hype of Generative AI need top-down regulation, or will it implode? Op. cit.
 83. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., . . . Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. doi:10.1038/sdata.2016.18.
 84. Mons, B. (2023). Does the hype of Generative AI need top-down regulation, or will it implode? Op. cit.
 85. Ibid.
 86. <https://leam.ai/feasibility-study-leam-2023/>.
 87. Cfr. <https://www.reuters.com/technology/microsofts-deal-with-mistral-ai-faces-eu-scrutiny-2024-02-27/>.
 88. Balland, P. & Renda, A. (2023). Forge ahead or fall behind. Why we need a United Europe of Artificial Intelligence, CEPS Explainer (https://cdn.ceps.eu/wp-content/uploads/2023/11/CEPS-Explainer-2023-13_United-Europe-of-Artificial-Intelligence.pdf).
 89. <https://www.appliedai.de/en/hub-en/ai-act-impact-survey>.
 90. COM(2024) 28 final, Brussels, 24.1.2024: <https://ec.europa.eu/newsroom/dae/redirection/document/101621>. Vedi anche: https://ec.europa.eu/commission/presscorner/detail/en/ip_24_383.
 91. Balland, P. & Renda, A. (2023). Andare avanti o restare indietro, op. cit., p. 5.
 92. https://www.bruegel.org/analysis/adapting-european-union-ai-act-deal-generative-artificial-intelligence#footnote4_30gc309.

Publisher

EIT Digital
Rue Guimard 7
1040 Brussels
Belgium
www.eitdigital.eu

Contact

info@eitdigital.eu

ISBN 978-91-87253-72-0



Co-funded by the
European Union