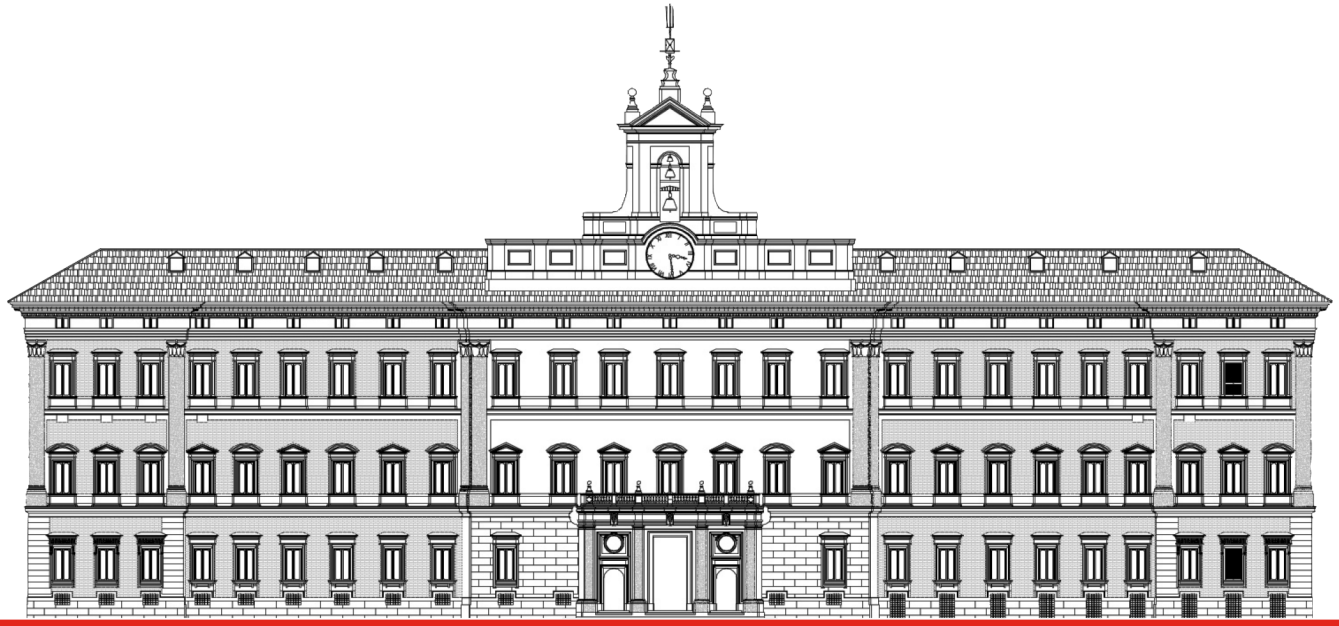
Camera dei deputati

XIX LEGISLATURA



Documentazione per le Commissioni

ATTIVITÀ DELL'UNIONE EUROPEA

Il regolamento UE in materia di intelligenza artificiale

n. 26

5 febbraio 2024

Camera dei deputati

XIX LEGISLATURA

Documentazione per le Commissioni
ATTIVITÀ DELL'UNIONE EUROPEA

Il regolamento UE in materia di intelligenza artificiale

n. 26

5 febbraio 2024

Il dossier è stato curato dall'**UFFICIO RAPPORTI CON L'UNIONE EUROPEA** (☏ 066760.2145 - ✉ *cdrue@camera.it* - 𝕏@CD_europa - europa.camera.it).

I N D I C E

IL REGOLAMENTO UE IN MATERIA DI INTELLIGENZA ARTIFICIALE

È in via di definitiva adozione da parte del Parlamento europeo e del Consiglio, secondo la **procedura legislativa ordinaria**, la proposta di regolamento, presentata dalla Commissione europea il 21 aprile 2021, recante un **quadro giuridico** in materia di **intelligenza artificiale** (esplicitamente denominato "**legge sull'intelligenza artificiale**").

In esito ai **triloghi** (negoziati interistituzionali tra i rappresentanti di Parlamento europeo, Consiglio e Commissione per concordare il testo da sottoporre all'approvazione dei due colegislatori), il **9 dicembre 2023**, al termine di una sessione negoziale di tre giorni, è stato infatti raggiunto un **accordo politico provvisorio**, con l'obiettivo di approvare in via definitiva la nuova normativa **entro la conclusione dell'attuale legislatura europea**.

L'accordo dovrà ora essere **formalmente approvato** dal Consiglio (a maggioranza qualificata), e dal Parlamento europeo (a maggioranza dei suoi componenti), al più tardi nella sessione del prossimo aprile. Il **2 febbraio** il Comitato dei rappresentanti permanenti degli Stati membri presso l'UE (**COREPER**), massimo organo preparatorio dei lavori del Consiglio**, ha approvato all'unanimità** il testo dell'accordo del 9 dicembre.

Prima dell'avvio dei negoziati, il **Consiglio dell'UE**, con un orientamento generale adottato all'unanimità il **6 dicembre 2022** e il **Parlamento europeo** con emendamenti approvati in plenaria (con 499 voti a favore, 28 contrari e 93 astensioni) il **14 giugno 2023** avevano definito le proprie posizioni.

Nel presente dossier si riporta una sintesi dei principali contenuti del testo risultante dall'accordo provvisorio.

Si illustra inoltre, in estrema sintesi, il pacchetto di proposte presentato il 24 gennaio 2024 della Commissione europea, a seguito all'accordo politico, al fine di sostenere le start-up e le PMI europee nello sviluppo dell'intelligenza artificiale.

Oggetto, definizione e ambito di applicazione

Il regolamento ha **l'obiettivo** di migliorare il funzionamento del mercato interno e promuovere l'**adozione di un'intelligenza artificiale affidabile e incentrata sull'uomo**, garantendo, nel contempo, un elevato livello di **protezione** della **salute**, della **sicurezza** e dei **diritti fondamentali** sanciti dalla Carta dell'UE, compresa la **democrazia,** lo **Stato di diritto** e la tutela dell'ambiente dagli effetti dannosi dei sistemi di intelligenza artificiale nell'Unione, nonché sostenendo l'innovazione.

La definzione di **sistema di IA** formulata nel regolamento è quella di “un sistema basato su una macchina progettato per funzionare con **diversi livelli di autonomia** e che può mostrare adattività dopo l'implementazione e che, per obiettivi espliciti o impliciti, deduce, dall'input che riceve, come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali”.

Alla Commissione europea è demandata l’adozione di **linee guida** sull’applicazione della richiamata definizione di sistema di IA.

La definizione originariamente proposta dalla Commissione è stata modificata nel corso del negoziato allo scopo di **allinearla** all'approccio dell'**OCSE**.

La nuova disciplina non intende introdurre una normazione esaustiva e dettagliata di ogni aspetto connesso all’IA, così definita. Tiene conto che diversi profili della materia sono già riconducibili ad **atti legislativi vigenti o in corso di adozione** a livello UE (*v. infra*). In particolare, il nuovo regolamento introduce:

a) regole armonizzate per **l'immissione sul mercato**, la messa in servizio e l'uso dei sistemi di intelligenza artificiale nell'Unione;

b) il **divieto** di determinate pratiche di intelligenza artificiale;

c) requisiti specifici per i **sistemi di IA ad alto rischio** e obblighi per gli operatori di tali sistemi;

d) regole di **trasparenza** armonizzate per alcuni sistemi di IA;

e) regole armonizzate specifiche per l'immissione sul mercato di modelli di **IA di uso generale**;

f) regole sul **monitoraggio** del mercato, sulla governance e sull'applicazione della **vigilanza** del mercato stesso;

g) misure a sostegno dell'**innovazione**, con particolare attenzione alle PMI, comprese le start-up.

Il regolamento si applicherà ai **soggetti pubblici e privati**, all'interno e all'esterno dell'UE, a condizione che il **sistema di IA** sia **immesso sul mercato dell'Unione** o che il suo uso abbia effetti su persone situate nell'UE.

Riguarderà sia i **fornitori** (ad es. uno sviluppatore di uno strumento di screening dei CV) che gli **operatori di sistemi di IA ad alto rischio** (ad es. una banca che acquista uno strumento di screening dei CV).

Gli **importatori** di sistemi di IA dovranno inoltre garantire che il fornitore straniero abbia già eseguito l'appropriata procedura di valutazione della conformità, che il sistema rechi una marcatura di conformità europea (CE) e sia corredato della documentazione e delle istruzioni per l'uso richieste.

Il regolamento prevede inoltre determinati **obblighi per i fornitori di modelli di IA per finalità generali,** compresi i **modelli di IA generativa** di grandi dimensioni. Invece, i fornitori di **modelli gratuiti e open source** saranno **esentati** dalla maggior parte di questi obblighi. Esenzione che tuttavia non riguarderà gli obblighi incombenti ai fornitori di modelli di IA per finalità generali che comportano **rischi sistemici**.

Il regolamento non **pregiudicherà le competenze degli Stati membri** in materia di **sicurezza nazionale**. Non si applicherà infatti ai sistemi di IA che sono esclusivamente per **scopi militari, di difesa o di sicurezza nazionale**, indipendentemente dal tipo di entità che svolge tali attività, nonché alle **attività di ricerca, sviluppo e prototipazione** che precedono l'immissione sul mercato o a persone che utilizzano l'IA per motivi non professionali.

Classificazione dei sistemi di IA e pratiche di IA vietate

Seguendo un **approccio "basato sul rischio"**, in base al quale tanto maggiore è il rischio, quanto più rigorose sono le regole, la nuova disciplina stabilisce obblighi per fornitori e operatori dei sistemi di IA a seconda del livello di rischio che l'IA può generare: i) un **rischio inaccettabile**; ii) un **rischio alto**; iii) un **rischio basso o minimo**. Sono stabiliti anche obblighi specifici per la trasparenza.

Pratiche di IA vietate per rischio inaccettabile

Saranno **vietati** i sistemi di IA che determinano un **rischio inaccettabile** per la sicurezza, i mezzi di sussistenza e i diritti delle persone. In questa categoria rientrano i sistemi che possono **manipolare il comportamento umano** come quelli che consentono di attribuire un "**punteggio sociale**" (*social scoring*), per finalità pubbliche e private, classificando le persone in base al loro comportamento sociale o alle loro caratteristiche personali, e **determinate applicazioni di polizia predittiva**. Saranno quindi **vietati**, in particolare:

- i sistemi di sfruttamento delle vulnerabilità delle persone e di **utilizzo di tecniche subliminali** ovvero deliberatamente **manipolative o ingannevoli**;
- i sistemi di **categorizzazione biometrica** delle persone fisiche sulla base di dati biometrici per dedurne o desumerne la razza, le opinioni politiche, l'appartenenza sindacale, le convinzioni religiose o filosofiche, la vita sessuale o l'orientamento sessuale (sarà ancora possibile filtrare set di dati basandosi su dati biometrici nel settore delle attività di contrasto);
- i sistemi di **identificazione biometrica in tempo reale in spazi accessibili al pubblico** (ossia il riconoscimento facciale mediante telecamere a circuito chiuso) da parte delle autorità di contrasto (con limitate **eccezioni**: *vedi oltre*);

- i **sistemi di riconoscimento delle emozioni utilizzati sul luogo di lavoro e negli istituti scolastici**, eccetto per motivi medici o di sicurezza (ad esempio il monitoraggio dei livelli di stanchezza di un pilota);
- l'**estrazione non mirata** (*scraping*) di immagini facciali da internet o telecamere a circuito chiuso per la creazione o l'espansione di banche dati;
- i sistemi che consentono di attribuire un "**punteggio sociale**" (*social scoring*), classificando o valutando le persone in base al loro comportamento sociale o alle loro caratteristiche personali.

Sistemi ad alto rischio

Il regolamento considera ad **alto rischio** un **numero limitato di sistemi di IA** che possono potenzialmente avere ripercussioni negative sulla sicurezza delle persone o sui loro diritti fondamentali (tutelati dalla **Carta dei diritti fondamentali dell'UE**).

Prima di **immettere un sistema di IA ad alto rischio sul mercato dell'UE**, o di farlo entrare in servizio, i fornitori dovranno sottoporlo a una **valutazione della conformità**. Dovranno, quindi, dimostrare che il loro sistema è **conforme ai requisiti obbligatori** per un'IA affidabile (ad esempio: qualità dei dati, documentazione e tracciabilità, trasparenza, sorveglianza umana, accuratezza, cibersicurezza e robustezza). Anche per i sistemi biometrici è sempre richiesta una valutazione della conformità da parte di terzi. La valutazione dovrà essere ripetuta in caso di modifica sostanziale del sistema o della sua finalità.

I sistemi di IA ad alto rischio dovranno essere **tecnicamente robusti** per garantire che la tecnologia sia adatta allo scopo e che i risultati falsi positivi/negativi non incidano in modo sproporzionato sui gruppi protetti (ad esempio, per origine razziale o etnica, sesso, età, ecc.). Dovranno, inoltre, essere **addestrati e testati con set di dati sufficientemente rappresentativi** per **ridurre al minimo il rischio di integrare distorsioni inique** nel modello e garantire che, se presenti, queste possano essere risolte mediante opportune misure di rilevazione, correzione e attenuazione. Dovranno anche essere **tracciabili e verificabili**, garantendo la **conservazione dell'opportuna documentazione**, compresi i dati utilizzati per addestrare l'algoritmo, fondamentali per le indagini *ex post*.

Si impone inoltre agli operatori che siano organismi di diritto pubblico o operatori privati che forniscono servizi pubblici, nonché agli operatori che forniscono sistemi ad alto rischio di effettuare una **valutazione d'impatto sui diritti fondamentali**.

La valutazione deve consistere in una descrizione dei processi dell'operatore in cui il sistema di IA ad alto rischio sarà utilizzato, del periodo di tempo e della frequenza in cui il sistema di IA ad alto rischio è destinato a essere utilizzato, delle categorie di persone fisiche e dei gruppi che possono essere interessati dal suo uso nel contesto specifico, dei rischi specifici di danno che possono incidere sulle categorie di persone o sui gruppi di

persone interessati, e in una descrizione dell'attuazione delle misure di sorveglianza umana e delle misure da adottare in caso di concretizzazione dei rischi.

I sistemi di IA che costituiscono componenti di sicurezza di prodotti disciplinati dalla **legislazione settoriale** dell'Unione saranno sempre considerati ad alto rischio, se soggetti a una valutazione della conformità da parte di terzi ai sensi della legislazione settoriale stessa. I fornitori di tali sistemi dovranno inoltre **attuare sistemi di gestione della qualità e del rischio** per garantire la conformità ai nuovi requisiti e ridurre al minimo i rischi per gli utenti e per le persone interessate, anche dopo l'immissione sul mercato di un prodotto.

I sistemi di IA ad alto rischio implementati da autorità pubbliche o entità che agiscono per loro conto dovranno essere **registrati in una banca dati pubblica dell'UE**. Ove tali sistemi non siano utilizzati per le attività di contrasto e relative al controllo della migrazione, dovranno essere registrati in una parte non pubblica della banca dati, che sarà accessibile solo alle autorità di controllo competenti.

Le **autorità di vigilanza** del mercato contribuiranno al monitoraggio successivo all'immissione sul mercato mediante audit e offrendo ai fornitori la possibilità di segnalare incidenti o violazioni gravi degli obblighi in materia di diritti fondamentali di cui sono venuti a conoscenza. Qualsiasi autorità di vigilanza del mercato potrà, per motivi eccezionali, autorizzare l'immissione sul mercato di una specifica IA ad alto rischio.

Tra i sistemi ad **alto rischio** rientrano, in particolare, quelli:

- di **identificazione biometrica remota**, categorizzazione biometrica e riconoscimento delle emozioni (al di fuori delle categorie vietate);
- utilizzati come componenti di sicurezza nella gestione e nel funzionamento delle **infrastrutture digitali critiche**, del **traffico stradale** e della **fornitura di acqua, gas, riscaldamento ed elettricità**;
- finalizzati a determinare **l'accesso, l'ammissione o l'assegnazione agli istituti di istruzione e formazione professionale** (ad esempio, per valutare i risultati dell'apprendimento e orientare il processo di apprendimento e il monitoraggio dei comportamenti disonesti);
- relativi alla **valutazione dell'occupazione**, ad ottimizzare la **gestione dei lavoratori** e l'**accesso al lavoro autonomo** (ad esempio, per pubblicare annunci di lavoro mirati, analizzare e filtrare le candidature e valutare i candidati);
- usati per determinare **l'accesso a servizi e a prestazioni pubblici e privati essenziali** (come, ad esempio, l'assistenza sanitaria);
- finalizzati alla **valutazione dell'affidabilità creditizia** delle persone fisiche, alla valutazione dei rischi finanziari, nonché alla determinazione dei prezzi in relazione ad assicurazioni sulla vita e assicurazioni sanitarie;

- utilizzati nelle attività di **contrasto**, di **gestione della migrazione, dell'asilo e del controllo delle frontiere, di amministrazione della giustizia,** nonché nello **svolgimento dei processi democratici** e per la valutazione e classificazione delle chiamate di emergenza.

Non sono invece inclusi i **sistemi di raccomandazione** delle **piattaforme online di dimensioni molto grandi** (utilizzati dalle aziende online per suggerire agli utenti prodotti, servizi o contenuti che potrebbero essere di loro interesse) in quanto sono **già disciplinati** da altre normative (regolamento sui mercati digitali e regolamento sui servizi digitali).

L'**elenco** dei sistemi di IA ad alto rischio, che può essere **modificato** per allineare la normativa **all'evoluzione tecnologica**, è **allegato** al regolamento.

Sistemi a rischio minimo

I sistemi di **IA a rischio minimo** (come **videogiochi o filtri spam)** saranno **esenti da obblighi**, ferma restando l'adesione volontaria a codici di condotta, da parte dei fornitori di tali sistemi, ad esempio laddove esista un evidente rischio di manipolazione. Gli utenti dovranno essere consapevoli del fatto che stanno interagendo con una macchina.

La **grande maggioranza dei sistemi di IA** attualmente utilizzati o il cui utilizzo è probabile nell'UE rientra in questa categoria.

Obblighi di trasparenza per determinati sistemi di IA e modelli di IA per finalità generali

A determinati sistemi di IA sono imposti **specifici obblighi di trasparenza**, ad esempio laddove esista un evidente rischio di manipolazione (come attraverso l'uso di **chatbot**); gli utenti dovranno essere consapevoli del fatto che stanno interagendo con una macchina. I fornitori di sistemi di IA, compresi i sistemi di IA per finalità generali (*General purpose AI - GPAI*), che generano contenuti audio, immagini, video o di testo sintetici, dovranno garantire che i risultati del sistema di IA siano contrassegnati in un formato leggibile dalla macchina e **rilevabili come generati o manipolati artificialmente**.

Anche i ***deep fake*** **dovranno essere etichettati come tali** e gli utenti dovranno essere informati quando vengono utilizzati sistemi di categorizzazione biometrica o di riconoscimento delle emozioni.

Il regolamento prende, quindi, in considerazione i **rischi sistemici** che potrebbero derivare dai **modelli di IA per finalità generali**, compresi i **modelli di IA generativa di grandi dimensioni** (*vedi oltre*), che possono essere utilizzati per un'ampia serie di compiti e stanno diventando la base di molti sistemi di IA nell'UE. Alcuni di questi modelli potrebbero comportare rischi sistemici se risultano

particolarmente efficaci o molto utilizzati. Modelli potenti potrebbero, ad esempio, causare incidenti gravi o essere utilizzati impropriamente per attacchi informatici di vasta portata.

Il "**rischio sistemico** a livello di Unione" si riferisce alla possibilità che l'uso dell'IA possa avere un **impatto significativo sul mercato** interno a causa della sua portata e con **effetti negativi reali o ragionevolmente prevedibili** su **salute pubblica**, **sicurezza, diritti fondamentali** o sulla **società nel suo insieme**, che possono essere propagati su larga scala lungo tutta la catena del valore. Ad esempio, se un'applicazione di guida autonoma mal funzionasse su larga scala, potrebbe causare incidenti stradali su vasta scala, influenzando quindi l'intero sistema di mobilità urbana.

Quanto al concetto di "**incidente grave**", ci si riferisce a qualsiasi incidente o malfunzionamento di un sistema di IA che porti direttamente o indirettamente a uno dei seguenti effetti: (a) la morte di una persona o un grave danno alla salute di una persona; b) un'interruzione grave e irreversibile della gestione e del funzionamento delle infrastrutture critiche; c) violazione degli obblighi derivanti dal diritto dell'Unione volti a tutelare i diritti fondamentali; d) danni gravi alla proprietà o all'ambiente.

Ad esempio, si pensi ad un errore in un sistema diagnostico medico basato sull'IA che porta a diagnosi errate e danni significativi ai pazienti.

L'Ufficio per l'IA (*v. infra*) incoraggia e facilita l'elaborazione di **codici di condotta** a livello di Unione per facilitare l'efficace attuazione degli obblighi in materia di rilevamento ed etichettatura di contenuti generati o manipolati artificialmente.

Eccezioni previste per le autorità di contrasto

L'uso dell'IA da parte delle forze dell'ordine ha rappresentato **uno dei punti più delicati e controversi del negoziato**. Il compromesso raggiunto consente il **riconoscimento biometrico da remoto** in tempo reale negli **spazi accessibili al pubblico** solo in alcuni casi quali:

1) **attività di contrasto** relative a **16 reati specifici**: terrorismo; tratta di esseri umani; sfruttamento sessuale di minori e materiale pedopornografico; traffico illecito di stupefacenti e sostanze psicotrope; traffico illecito di armi, munizioni ed esplosivi; omicidio volontario; lesioni personali gravi; traffico illecito di organi e tessuti umani; traffico illecito di materie nucleari e radioattive, rapimento, sequestro e presa di ostaggi; reati che rientrano nella competenza giurisdizionale della Corte penale internazionale; dirottamento di un aeromobile o una nave; stupro; reati ambientali; furto organizzato o rapina a mano armata; sabotaggio, partecipazione a un'organizzazione criminale coinvolta in uno o più dei reati elencati sopra;

2) **ricerca mirata** di specifiche vittime, rapimento, tratta e sfruttamento sessuale di esseri umani e persone scomparse;

3) **prevenzione di minacce** per la vita o l'incolumità fisica delle persone o risposta da una minaccia attuale o prevedibile di **attacco terroristico**.

Si specifica che **l'identificazione biometrica può assumere varie forme**. Può essere utilizzata per l'autenticazione degli utenti, ad esempio per sbloccare uno *smartphone* o per le verifiche e l'autenticazione presso i valichi di frontiera nei controlli dell'identità e dei documenti di viaggio di una persona (corrispondenza "uno a uno"). L'identificazione biometrica remota potrebbe anche essere utilizzata per identificare persone nella folla, ad esempio confrontando l'immagine di una persona con quelle contenute in una banca dati (corrispondenza "uno a molti").

L'accuratezza dei sistemi per il **riconoscimento facciale** può variare in modo significativo in base a un'ampia gamma di fattori, quali la qualità della fotocamera, la luce, la distanza, la banca dati, l'algoritmo e l'etnia, l'età o il sesso del soggetto. Lo stesso vale per il riconoscimento vocale e dell'andatura e per altri sistemi biometrici. Il tasso di falsi positivi dei sistemi altamente avanzati è in continua diminuzione. Un tasso di accuratezza del 99% può sembrare adeguato, mentre è notevolmente rischioso quando può condurre a sospettare di una persona innocente. D'altra parte, anche un tasso di errore dello 0,1% è molto elevato se riguarda decine di migliaia di persone.

L'identificazione biometrica remota in tempo reale da parte delle autorità di contrasto sarà subordinata a un'**autorizzazione preventiva rilasciata da un'autorità giudiziaria o amministrativa indipendente**. In caso di **urgenza** debitamente giustificata, tuttavia, si potrà procedere **senza un'autorizzazione**, purché quest'ultima sia richiesta senza indebito ritardo, al più tardi entro 24 ore; se l'autorizzazione **non è concessa**, è necessario che tutti i dati e gli output siano soppressi. L'autorizzazione dovrà essere preceduta da una **valutazione preventiva d'impatto sui diritti fondamentali** e dovrà essere **notificata all'autorità di vigilanza e all'autorità per la protezione dei dati interessate**.

L'uso di sistemi di IA per l'**identificazione biometrica remota a posteriori** delle persone oggetto di indagine (identificazione di persone in materiale video raccolto in precedenza) richiederà l'autorizzazione preventiva di un'autorità giudiziaria o di un'autorità amministrativa indipendente e la notifica all'autorità per la protezione dei dati e all'autorità di vigilanza del mercato.

Sistemi di IA per finalità generali e modelli fondativi o di base

La nuova disciplina, introdotta nell'accordo di compromesso finale, su richiesta del PE, reca disposizioni dettagliate applicabili alle situazioni in cui i sistemi di IA per finalità generali *(general purpose AI*) sono utilizzati per scopi per i quali non sono stati **intenzionalmente e specificamente progettati** e in cui la tecnologia viene successivamente integrata in un altro sistema ad alto rischio. I modelli di IA per finalità generali **comprendono i modelli di IA generativa di grandi dimensioni**.

È considerato importante che un fornitore che intenda basarsi su un modello di IA per finalità generali disponga di tutte le informazioni necessarie per far sì che il suo sistema sia sicuro e conforme al regolamento. Il regolamento obbliga i fornitori di tali modelli a **comunicare determinate informazioni ai fornitori di sistemi a valle**. I fornitori dovranno inoltre disporre di politiche in essere atte a garantire il **rispetto del diritto d'autore** nel corso della formazione dei loro modelli.

Alcuni di questi modelli potrebbero comportare **rischi sistemici** dato che sono particolarmente efficaci o molto utilizzati.

La Commissione europea ritiene che, allo stato attuale della tecnologia, i modelli di IA per finalità generali che sono stati addestrati utilizzando **una potenza di calcolo totale superiore a 10^25 FLOPS** comportino **rischi sistemici**.

La soglia stabilita è superiore a quanto inizialmente voluto dal Parlamento europeo, ma inferiore a quella prevista dal mandato del Consiglio, che era di 10^26 FLOPS.

L'Ufficio per l'IA (istituito all'interno della Commissione) potrà **aggiornare tale soglia alla luce dell'evoluzione tecnologica** ovvero integrare il criterio della potenza di calcolo con altri criteri (ad esempio, il numero di utenti o il grado di autonomia del modello).

I fornitori di modelli che comportano rischi sistemici saranno pertanto tenuti a **valutare e attenuare i rischi**, a **segnalare incidenti gravi**, a **condurre prove e valutazioni dei modelli all'avanguardia**, a garantire la **cibersicurezza** e a fornire **informazioni sul consumo energetico** dei loro modelli. A tal fine potranno **collaborare con l'Ufficio europeo per l'IA** per elaborare codici di condotta. Un **gruppo di esperti scientifici** svolgerà un ruolo centrale nella supervisione dei modelli di IA per finalità generali.

La Commissione spiega per quale motivo 10^25 FLOPS è una soglia adeguata per ritenere che un'IA per finalità generali possa comportare rischi sistemici. La soglia, infatti, riflette i modelli di IA per finalità generali attualmente più avanzati, ossia **GPT-4 di OpenAI** e probabilmente **Gemini di Google DeepMind**. Non essendo ancora sufficientemente comprese le capacità dei modelli al di sopra di tale soglia, si ritiene che essi possano comportare rischi sistemici e che sia ragionevole, pertanto, imporre ai fornitori obblighi aggiuntivi.

Sono state inoltre concordate **regole specifiche per i modelli fondativi o di base** (la cui regolamentazione non era presente nella proposta originaria e che ha rappresentato una **tematica particolarmente divisiva nel negoziato**): i grandi sistemi in grado di svolgere con competenza un'ampia gamma di compiti distintivi, quali la generazione di video, testi, immagini, il calcolo di dati o la generazione di codici informatici. I modelli di base debbano **rispettare specifici obblighi di trasparenza** prima di essere immessi sul mercato. È stato introdotto un regime più rigoroso per i modelli di base "**ad alto impatto**", come **GPT-4,** per i quali occorre una verifica prima dell'immissione nel mercato con riguardo alla sicurezza informatica e alla trasparenza, nonché una condivisione della documentazione

tecnica. Si tratta di modelli di base addestrati con grandi quantità di dati e di complessità, capacità e prestazioni avanzate ben al di sopra della media, che possono diffondere i rischi sistemici lungo la catena del valore.

Misure a sostegno dell'innovazione

Il regolamento contiene diverse **misure a sostegno dell'innovazione**. In particolare, consente la creazione di **spazi di sperimentazione normativa per l'IA** (*sandbo*x normativi) e di **prova in condizioni reali**, che forniscono un ambiente controllato per testare tecnologie innovative per un periodo di tempo limitato, promuovendo in tal modo l'innovazione da parte delle imprese, delle PMI e delle *start-up*.

Gli Stati membri dovranno istituire **almeno un sandbox normativo sull'IA a livello nazionale**. Potrà anche essere istituito **congiuntamente** tra più Stati membri.

Le **prove in condizioni reali dei sistemi di IA ad alto rischio** potranno essere effettuate per un massimo di 6 mesi (prorogabili di altri 6 mesi). Prima delle prove dovrà essere elaborato un piano da presentare all'autorità di vigilanza del mercato, la quale dovrà approvare il piano e le condizioni di prova specifiche; in caso di mancata risposta entro 30 giorni, il piano si considererà tacitamente approvato. Le prove potranno essere oggetto di ispezioni senza preavviso da parte dell'Autorità.

Le prove in condizioni reali potranno essere effettuate solo se sono presenti garanzie specifiche: ad esempio, gli utenti dei sistemi sottoposti a prova in condizioni reali dovranno fornire un consenso informato, le prove non dovranno avere alcun effetto negativo sugli utenti, gli esiti delle prove dovranno essere reversibili o poter essere ignorati, i relativi dato dovranno essere cancellati dopo la conclusione delle prove. Una protezione speciale dovrà essere riservata ai gruppi vulnerabili, ad esempio a causa della loro età o della disabilità fisica o mentale.

Architettura di governance

Le **autorità nazionali competenti per la vigilanza del mercato** sorveglieranno l'attuazione delle nuove norme a livello nazionale, mentre un **Ufficio europeo per l'IA,** costituito presso la Commissione europea, garantirà il coordinamento a livello europeo. Ciascuno Stato membro designerà una o più autorità nazionali competenti, incaricate di supervisionarne l'applicazione e l'attuazione, nonché di svolgere attività di vigilanza del mercato.

Un **comitato scientifico di esperti indipendenti** avrà il compito di segnalare i rischi sistemici e contribuire alla classificazione e alla sperimentazione dei modelli. Un **comitato europeo per l'IA**, composto dai rappresentanti degli Stati membri, svolgerà il ruolo di **piattaforma di coordinamento e di organo consultivo** per la

Commissione europea. Il Garante europeo della protezione dei dati parteciperà come osservatore.

Si prevede infine l'istituzione di un **forum consultivo** per i portatori di interessi, come i rappresentanti dell'industria, le PMI, le *start-up*, la società civile e il mondo accademico.

Più nello specifico, la **missione dell'Ufficio per l'IA** sarà quella di sviluppare le competenze e le capacità dell'Unione nel settore dell'IA e contribuire all'attuazione della legislazione dell'Unione in materia di intelligenza artificiale in una struttura centralizzata. In particolare, l'Ufficio avrà il compito di **applicare e supervisionare le nuove regole per i modelli di IA per finalità generali,** con il potere di richiedere documentazione, condurre valutazioni dei modelli, indagare sulle segnalazioni e chiedere ai fornitori di adottare misure correttive.

L'Ufficio garantirà inoltre il **coordinamento** per quanto riguarda la politica in materia di IA e la collaborazione tra le istituzioni, gli organi e le agenzie dell'Unione coinvolti, nonché con gli esperti e i portatori di interessi. In particolare, dovrà provvedere a creare un forte legame con la comunità scientifica per sostenere l'applicazione delle regole, fungerà da punto di riferimento internazionale per gli esperti e le organizzazioni di esperti indipendenti e faciliterà gli scambi e la collaborazione con istituzioni analoghe in tutto il mondo.

Sistemi di IA già immessi sul mercato o messi in servizio

Per quanto riguarda i **sistemi di IA già immessi sul mercato o messi in servizio**, l'accordo di compromesso stabilisce in particolare che: le autorità pubbliche che sono fornitori o utilizzatori di sistemi di IA ad alto rischio avranno **quattro anni** di tempo per rendere conformi i propri sistemi; i modelli GPAI immessi sul mercato prima della data di applicazione delle disposizioni a loro relative (ossia 12 mesi dopo l'entrata in vigore del regolamento) avranno **due anni** di tempo dalla data di applicazione di tali disposizioni (quindi 3 anni in totale) per conformarsi.

Sanzioni

Le **sanzioni** per le violazioni del nuovo regolamento sono state fissate in una **percentuale del fatturato annuo globale** nell'anno finanziario precedente della società incriminata **o in un importo predeterminato**, a seconda di quale sia il più elevato.

Più nello specifico, per i sistemi di IA che sono immessi sul mercato o messi in servizio e che non rispettano i requisiti del regolamento, gli Stati membri dovranno stabilire **sanzioni** effettive, proporzionate e dissuasive, comprese sanzioni amministrative pecuniarie, in relazione alle violazioni, e comunicarle alla Commissione. Il regolamento stabilisce le **soglie** da tenere in considerazione:

- fino a **35 milioni di euro o al 7% del fatturato** mondiale totale annuo dell'esercizio precedente (se superiore) per violazioni relative a pratiche vietate o per l'inosservanza di requisiti in materia di dati;
- fino a **15 milioni di euro o al 3% del fatturato mondiale** totale annuo dell'esercizio precedente, per l'inosservanza di qualsiasi altro requisito o obbligo del regolamento, compresa la violazione delle regole relative ai modelli di IA per finalità generali;
- fino a **7,5 milioni di euro o all'1,5% del fatturato mondiale** totale annuo dell'esercizio precedente, per la fornitura di informazioni inesatte, incomplete o fuorvianti agli organismi notificati e alle autorità nazionali competenti in risposta a una richiesta;
- per ciascuna categoria di violazione, la **soglia per le PMI** sarebbe l'importo più basso tra i due previsti, mentre per le altre imprese sarebbe l'importo più elevato.

La Commissione, sulla base del parere del comitato, elaborerà orientamenti al fine di armonizzare le regole e le prassi nazionali in materia di calcolo delle sanzioni amministrative pecuniarie.

Le istituzioni, le agenzie o gli organismi dell'UE non beneficeranno di deroghe; il Garante europeo della protezione dei dati avrà il potere di infliggere loro sanzioni pecuniarie.

L'accordo di compromesso chiarisce inoltre che una persona fisica o giuridica potrà presentare un **reclamo** alla pertinente **autorità di vigilanza del mercato** riguardo alla non conformità con il regolamento sull'IA e potrà attendersi che tale reclamo sia trattato in linea con le procedure specifiche di tale autorità.

Misure di attuazione

Il regolamento prevede un ampio ricorso a misure di attuazione, adottate dalla Commissione europea, al fine sia di definire aspetti di dettaglio della disciplina sia, soprattutto, di aggiornala in base all’evoluzione tecnologica e di contesto.

A questo scopo diverse disposizioni del regolamento potranno essere modificate o integrate mediante **atti delegati e di esecuzione** emessi dalla **Commissione europea** secondo le procedure disciplinate dal medesimo regolamento in coerenza con gli articoli **290 e 291 del TFUE.**

Si prevede il ricorso agli atti delegati, tra le altre cose, per aggiornare la soglia FLOPS e aggiungere criteri per classificare i modelli di IA per finalità generali come modelli che presentano rischi sistemici.

Con atti esecutivi potranno invece essere, tra l’altro, modificate le modalità per istituire spazi di sperimentazione normativa ed elementi del piano di prova in condizioni reali.

Il regolamento demanda inoltre alla Commissione l'elaborazione di **linee guida sull'attuazione pratica** di diversi aspetti della disciplina, relativi in particolare a: a) l'applicazione dei requisiti e degli obblighi per i sistemi ad alto rischio e in materia di responsabilità lungo la catena del valore dell'IA; b) le pratiche di IA vietate; c) l'attuazione pratica delle disposizioni relative alla modifica sostanziale; d) l'attuazione pratica degli obblighi di trasparenza; e) i rapporti tra il regolamento e altre determinate legislazioni e pertinenti normative dell'Unione, anche per quanto riguarda la coerenza della loro applicazione; f) l'applicazione della definizione di sistema di IA.

Nel pubblicare le linee guida, la Commissione deve prestare particolare attenzione alle esigenze delle PMI, comprese le start-up, delle autorità pubbliche locali e dei settori maggiormente interessati dal regolamento. Le linee guida devono tenere debitamente conto tra l'altro dello stato dell'arte generale delle conoscenze in materia di IA. Su richiesta degli **Stati membri o dell'Ufficio IA,** o di propria iniziativa, la Commissione aggiorna le linee guida ove ritenuto necessario.

Entrata in vigore

Il regolamento sull'IA sarà pienamente applicabile **due anni** dopo la sua entrata in vigore, secondo il seguente approccio graduale:

- **6 mesi** dopo l'entrata in vigore, gli Stati membri devono eliminare gradualmente i sistemi vietati;
- **12 mesi** dopo: diventano applicabili gli obblighi relativi alla *governance* dell'IA per finalità generali;
- **24 mesi** dopo: tutte le regole della legge sull'IA diventano applicabili, compresi gli obblighi per i sistemi ad alto rischio definiti nell'allegato III (elenco dei casi d'uso ad alto rischio);
- **36 mesi** dopo: si applicano gli obblighi per i sistemi ad alto rischio definiti nell'allegato II (elenco della normativa di armonizzazione dell'Unione).

La posizione negoziale del Governo italiano

Nel corso del complesso negoziato in seno al Consiglio, il **Governo italiano** si è sempre dichiarato a **favore** dell'introduzione di un quadro comune di regole sull'intelligenza artificiale, sottolineando l'importanza che il nuovo regolamento tutelasse i **diritti fondamentali**, imponesse **obblighi e sanzioni commisurati al rischio** e allo stesso tempo permettesse di **mantenere il passo tecnologico e lo slancio verso l'innovazione** di altri competitor globali, come Stati Uniti e Cina.

La **posizione negoziale italiana**, inoltre, si è basata su una **visione "umano-centrica"**, volta a promuovere pertanto la **semplificazione** e la **chiarezza** delle **definizioni** a partire da quelle di **IA, sistemi generativi, *foundation model*.**

Sulla base di questa impostazione, il Governo italiano ha insistito per la **compartimentazione di alcuni perimetri e settori esclusi** dall'ambito di applicazione della originaria proposta di regolamento, oltre che sulla maggior chiarezza rispetto al coordinamento con le normative di settore, in particolare **bancario e assicurativo**.

Si è espresso inoltre a favore di un sistema di ***self-assessment*** **da parte delle aziende dei sistemi di IA**, attraverso linee guida o un archivio di esempi. Quanto agli **obblighi**, ha supportato la definizione di oneri e obblighi lungo la catena del valore dei sistemi IA e insistito affinché quest'ultimi non risultino **troppo gravosi soprattutto per le PMI**.

Sempre per quanto concerne le **definizioni**, l'Italia ha chiesto chiarezza relativamente ai **sistemi a finalità generale** in quanto vi è il **rischio di rapida obsolescenza di definizioni troppo ristrette**.

Ha altresì espresso apertura rispetto alla possibilità per gli Stati membri di integrarsi attraverso **sandbox a livello unionale**. Ha chiesto pertanto di estendere la **presunzione di conformità** con la normativa UE anche ai **risultati** delle sandbox, per accelerare il processo di integrazione dei medesimi risultati nel contesto normativo.

L'Italia ha manifestato anche **apertura con riferimento alle valutazioni di impatto sui diritti fondamentali**, in linea con le richieste del Parlamento europeo e nazionale di maggiore attenzione ai profili etici.

Infine, il nostro Paese ha chiesto di includere tra i sistemi ad alto rischio i sistemi di ausilio e supporto al giudice nella ricerca e nell'interpretazione dei fatti e della legge, nonché i sistemi di ADR (metodi alternativi di risoluzione delle controversie).

Per definire la convergenza su alcuni dei punti più complessi del negoziato, **Italia, Francia e Germania** il 30 ottobre scorso avevano tenuto un incontro trilaterale a Roma (comunicato stampa) durante il quale sono state discusse soprattutto le tematiche relative alla regolamentazione dei modelli fondativi, l'esclusione delle forze dell'ordine e della sicurezza nazionale e la classificazione dei sistemi ad alto rischio.

In esito all'incontro, i tre Paesi avevano sottoscritto un **documento di posizione** congiunto che si opponeva, con riferimento ai **modelli di fondazione**, all'**approccio graduale** della Presidenza spagnola basato su criteri e regole più rigorosi per i modelli ad alto impatto. I tre paesi avevano invece **chiesto** che nella prima fase i modelli in questione siano disciplinati da un **codice di condotta** e che le future sanzioni europee non vengano imposte in prima istanza.

In sostanza, i tre paesi hanno chiesto di **non applicare regole troppo stringenti ai sistemi di intelligenza artificiale generativa**, ma che le imprese possano sviluppare i modelli di fondazione, ossia i sistemi alla base dei prodotti di IA, come ChatGPT, semplicemente autocertificandone le caratteristiche e adeguandosi alle migliori pratiche.

Pacchetto per l'innovazione in materia di IA a sostegno delle *start-up* e delle PMI nel settore dell'intelligenza artificiale

Facendo seguito all'accordo politico, il **24 gennaio 2024** la Commissione europea ha varato un **pacchetto di misure** volto a **sostenere le *start-up* e le PMI europee nello sviluppo dell'intelligenza artificiale**. Il pacchetto:

- la proposta di regolamento che modifica il regolamento (UE) 2021/1173 relativo all'istituzione dell'impresa comune per il calcolo ad alte prestazioni europeo (**EuroHPC**) per istituire **fabbriche di IA**, (impresa comune dei supercomputer dell'UE);
- la **decisione istitutiva del richiamato Ufficio per l'IA** in seno alla Commissione;
- la **comunicazione** sulle **start-up e l'innovazione in materia di IA** che delinea ulteriori **attività chiave**, tra cui:

  a) il sostegno finanziario attraverso Orizzonte Europa e il programma Europa digitale dedicato all'IA generativa (che dovrebbe generare un ulteriore investimento pubblico e privato complessivo di circa **4 miliardi di euro fino al 2027**);

  b) alcune iniziative volte a **rafforzare il bacino generativo di talenti** dell'UE in materia di IA attraverso attività di istruzione, formazione, qualificazione e riqualificazione;

  c) la promozione di investimenti pubblici e privati nelle start-up e nelle scale-up nel settore dell'IA, anche attraverso **il capitale di rischio o il sostegno** al **capitale** (ad es. nuove iniziative del programma di accelerazione del CEI e di InvestEU);

  d) l'accelerazione dello **sviluppo e della diffusione di** spazi **comuni europei di dati,** messi a disposizione della comunità dell'IA, per i quali i dati sono una risorsa fondamentale per formare e migliorare i loro modelli;

  e) l'iniziativa **"GenAI4EU"**, che mira a sostenere lo sviluppo di nuovi casi d'uso e applicazioni emergenti nei 14 ecosistemi industriali europei, nonché nel settore pubblico. I settori di applicazione comprendono la robotica, la salute, le biotecnologie, l'industria manifatturiera, la mobilità, il clima e i mondi virtuali.

La Commissione ha annunciato inoltre che sta istituendo, con diversi Stati membri, **due consorzi** per **l'infrastruttura digitale europea (EDIC)**, ovvero l'**"Alleanza per le tecnologie linguistiche" (ALT-EDIC)**, che mira a sviluppare un'infrastruttura europea comune nelle tecnologie del linguaggio, volta anche a sostenere lo sviluppo di grandi modelli linguistici europei, e l' **EDIC "CitiVERSE"**, che tra l'altro aiuterà le città a simulare e ottimizzare i processi, dalla gestione del traffico alla gestione dei rifiuti. Gli Stati membri istituiranno ora i consorzi per l'infrastruttura digitale europea ALT-EDIC e CitiVERSE EDIC con il sostegno della Commissione.

La Commissione ha adottato anche una comunicazione che delinea **l'approccio strategico sull'uso dell'intelligenza artificiale**, anticipando e preparandosi all'attuazione del regolamento sull'IA. Comprende azioni concrete per garantire lo sviluppo e l'uso di un'IA affidabile, sicura ed etica. La Commissione annuncia altresì che si sta preparando a sostenere le PA dell'UE nell'adozione e nell'utilizzo dell'intelligenza artificiale.