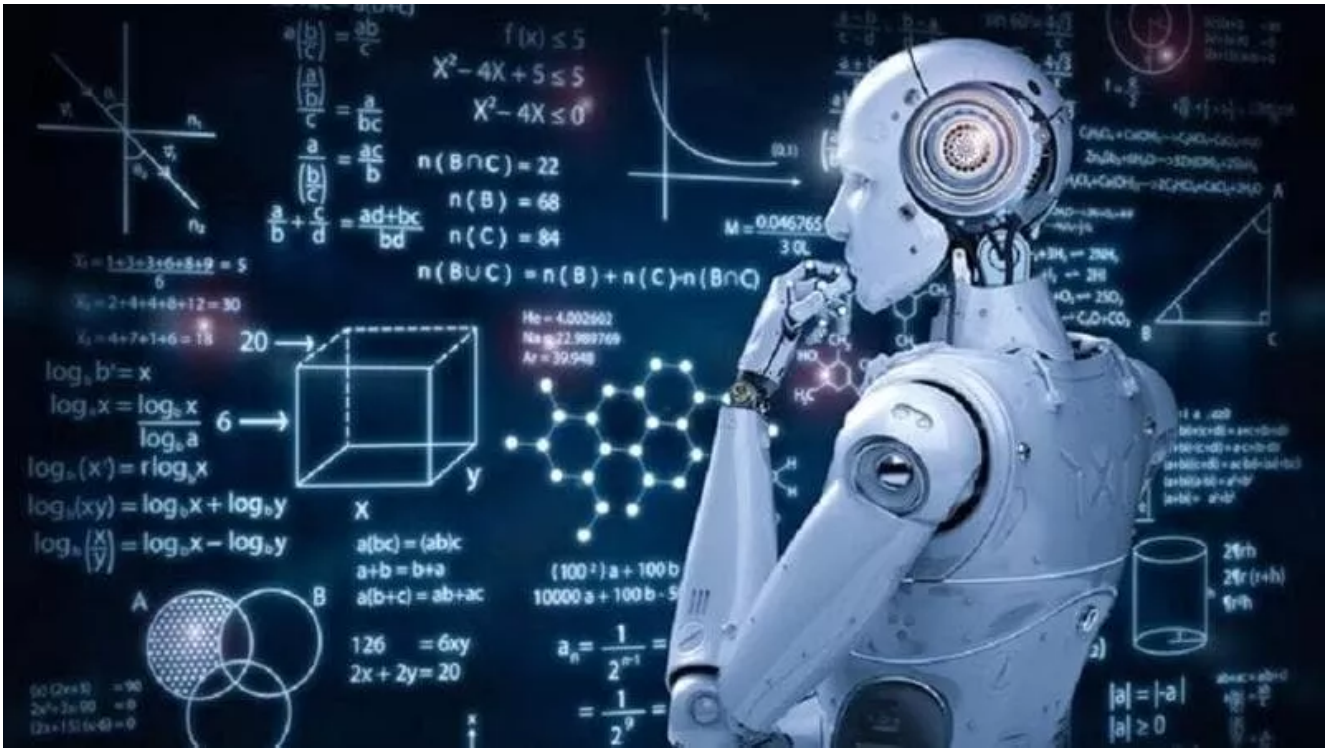
Ai Act, sulle regole per l’intelligenza artificiale raggiunto l’accordo in Europa. Breton: “Momento storico, guideremo la corsa”



L’Unione europea avrà una pionieristica legislazione sull’**Intelligenza artificiale**, la più completa ed organica al mondo, con l’obiettivo di coniugarne lo sviluppo con il rispetto dei diritti fondamentali. Sabato notte – dopo 36 ore di negoziato finale condotto nell’arco di tre giorni – l’Europarlamento, la Commissione e il Consiglio hanno trovato un accordo politico sull’**AI Act**. Il testo finale andrà ancora limato nelle prossime settimane, ma l’intesa assicura che sarà approvato entro la fine della legislatura europea, per poi entrare progressivamente in vigore nei successivi due anni. Un risultato non scontato, considerate le distanze con cui da un lato il Parlamento, più attento alla protezione dei diritti, e

dall’altro i governi, più attenti alle ragioni dello sviluppo economico e dell’ordine pubblico, si erano presentati a questo ultimo appuntamento negoziale. Lo scoglio sulla regolamentazione delle Intelligenze artificiali più potenti, come quelle sviluppate dai colossi **OpenAi, Meta e Google**, era stata superato giovedì: saranno sottoposte a regole vincolanti su trasparenza e sicurezza, non solo a codici di condotta volontari. Ieri poi si è trovato un compromesso anche sull’ultimo, l’utilizzo delle applicazioni di AI nei contesti di polizia: resta ammesso il riconoscimento biometrico, per citare uno dei temi più spinosi, ma solo in caso di reati gravi e previa autorizzazione di un giudice.

“E’ un momento storico”, ha esultato il Commissario europeo **Thierry Breton**, definendo l’AI Act una rampa di lancio che permetterà a ricercatori e aziende europee “di guidare la corsa globale all’AI”. “Questa legge assicura che i diritti e le libertà siano al centro dello sviluppo di questa tecnologia rivoluzionaria, garantendo un bilanciamento tra innovazione e protezione”, ha detto **Brando Benifei**, europarlamentare Pd e relatore della norma. Anche se le incognite sull’equilibrio finale, l’applicazione e l’efficacia dell’AI Act restano molte.

Emozioni, discriminazioni e supervisione umana

Il principio alla base dell’AI Act è la distinzione delle applicazioni dell’Intelligenza artificiale sulla base del livello di rischio che pongono per i diritti fondamentali. Una serie di ambiti giudicati troppo rischiosi sono quindi banditi: si tratta per esempio dei sistemi di **“social scoring”** (come quello teorizzato e sperimentato in Cina), o di quelli che manipolano comportamenti e decisioni. Gli algoritmi di **riconoscimento delle emozioni** vengono banditi da scuole e luoghi di lavoro, ma sembrerebbe restino utilizzabili in contesti di immigrazione e sicurezza, come chiedevano i

governi. La spunta invece il Parlamento sui sistemi di categorizzazione basati su informazioni sensibili – razza, religione, orientamento sessuale -: saranno vietati.

C'è poi una lunga serie di applicazioni giudicate ad **"alto rischio"** e sono quelle che riguardano ambiti che toccano i diritti fondamentali come salute, lavoro, educazione, immigrazione, giustizia. Qui l'AI Act introduce una serie di prescrizioni per chi le sviluppa e per chi le utilizza, come per esempio una valutazione preliminare dell'impatto, anche per evitare i rischi – ben documentato – di errori o discriminazioni, la necessità di una **supervisione umana**, quella di informare l'utilizzatore che sta interagendo con una macchina.

Riconoscimento facciale ed eccezioni per la polizia

Come detto, un punto molto dibattuto è stato l'utilizzo dei **sistemi di riconoscimento biometrico**. Il Parlamento – con l'appoggio di diverse organizzazioni per i diritti civili – aveva proposto un bando completo, mentre i governi volevano ampissime eccezioni per i contesti di sicurezza. Il compromesso è che potranno essere usati solo previa autorizzazione di un giudice e in circostanze ben definite. Quelli "ex post", quindi su immagini registrate solo per cercare persone sospettate di crimini gravi, quelli in tempo reale solo per emergenze terroristiche, ricerca di vittime o di sospettati di crimini gravi. Le eccezioni per la polizia riguardano anche l'utilizzo di applicazioni alto rischio, che potranno essere impiegate anche prima di aver ricevuto l'attestazione di conformità su autorizzazione di un giudice.

La legge non pone invece alcune limite all'utilizzo degli algoritmi nell'ambito della difesa e militare, che è esclusiva competenza degli Stati membri.

ChatGPT e i suoi fratelli

Rispetto al testo originario della Commissione, che risale a due anni fa, questo accordo aggiunge una serie di prescrizioni per le cosiddette General purpose AI, cioè grandi modelli così potenti da prestarsi a molteplici utilizzi. E' il caso di quello alla base di ChatGPT e di quelli sviluppati dagli altri big della Silicon Valley come Google o Meta. Queste prescrizioni saranno vincolanti: una vittoria del Parlamento, visto che i governi – in particolare Germania, Francia e Italia – avevano chiesto nei giorni scorsi di limitarsi a dei semplici codici di condotta, nel timore che una regolazione troppo stringente finisca per soffocare l'innovazione in Europa.

La norma su questi grandi modelli ha due livelli. Il primo, che si applica a tutti, prevede la pubblicazione di una lista dei materiali usati per l'addestramento degli algoritmi, strumento che in teoria dovrebbe aiutare i produttori di contenuti a difendere – o farsi riconoscere – i diritti d'autore, oltre all'obbligo di rendere riconoscibili – per contrastare truffe o disinformazione – tutti i contenuti prodotti all'AI. Il secondo livello si applicherà invece ai sistemi più potenti, quelli che pongono **"rischi sistemici"**, e prevede delle valutazioni di questi pericoli e delle strategie di mitigazione, oltre che l'obbligo di comunicare alla Commissione, che si doterà di un apposito AI Office, eventuali incidenti. Il mancato rispetto delle regole comporta multe che vanno dall'1,5 al 7% del fatturato globale delle aziende coinvolte.

Le incognite

Le disposizioni dell'AI Act, una volta approvato, entreranno in vigore in maniera progressiva: dopo sei mesi quelle sulle applicazioni proibite, dopo dodici quelle sui sistemi ad alto rischio e sui modelli più potenti, le ultime **dopo due anni**.

Sono tempi che serviranno alla Commissione per stabilire i dettagli tecnici necessari all'implementazione e alle aziende per adattarsi, anche se nel frattempo saranno già incoraggiate ad adeguarsi volontariamente. Ma certo questo lungo periodo di avvio alimenta i dubbi di chi sostiene che per una legge sarà difficilissimo tenere il passo di una tecnologia che – grazie all'abilità degli ingegneri e ai miliardi investiti – evolve in modo esponenziale. E che più la normativa è dettagliata più rischia di essere inefficace. L'alternativa, d'altra parte, è non prevedere alcuna regola o affidarsi ai codici di condotta autonomamente elaborati dalle stesse aziende: un approccio che è quello adottato per ora in sede di G7 e anche negli Stati Uniti, ma che consegna a Big Tech il potere di autoregolarsi. L'Europa pensa che non basti.

L'altra incognita riguarda la possibilità che questa normativa finisca per danneggiare l'innovazione in Europa più che favorirla. L'idea, anche questa diffusa, è che l'arbitro – in questo caso la Ue – non vince mai. Ma i rischi posti dall'Intelligenza artificiale, enormi come le opportunità, fanno dire a tanti altri che un arbitro in questo caso è necessario. Con la speranza che un campo da gioco delimitato con chiarezza aiuti a far crescere fuoriclasse in grado di competere con quelli americani.