AI Act, che cosa è il «rischio inaccettabile?». E cosa prevede la legge europea sull'intelligenza artificiale?

I quattro livelli di rischio, la regolamentazione dell'AI generativa, le applicazioni proibite, le sanzioni: cosa sapere sul nuovo regolamento approvato dal Parlamento europeo
(Fonte: https://www.corriere.it/ 14 marzo 2024)

AI Act, il lungo iter legislativo

Dopo quattro anni di bozze, revisioni, proposte e analisi, **il Parlamento europeo** in seduta plenaria ha dato la sua approvazione, con 523 voti a favore, all'**AI Act** (qui il testo)**.** Manca solo il voto - a questo punto formale - del Consiglio europeo e poi **la prima legge al mondo** per provare a regolamentare l'intelligenza artificiale. Il lungo iter legislativo è iniziato nel **2021** e terminerà verosimilmente **entro giugno**: l'obiettivo è quello di pubblicare la legge nella Gazzetta Ufficiale dell'Unione europea entro la fine della legislatura del Parlamento - il voto per le prossime elezioni è previsto per l'8 e il 9 giugno. Venti giorni dopo, entrerà ufficialmente in vigore, ma la sua applicazione concreta avrà tempi variabili a seconda dell'urgenza delle norme specifiche (ne parliamo nel dettaglio nelle prossime schede). L'obiettivo dell'AI Act lo ha riassunto l'eurodeputato **Brando Beniferi**, relatore della legge: «Rappresenta un chiaro percorso per lo **sviluppo sicuro e umanocentrico** dell'intelligenza artificiale». L'Ue si conferma ancora una volta **l'ente istituzionale più attivo nella ricerca di regole e buone norme sullo sviluppo della tecnologia**. Un percorso iniziato nel 2018, con il Gdpr, e proseguito con Digital Markets Act, Digital Services Act, Data Act e infine l'AI Act. La legge, questa, **più complessa da**

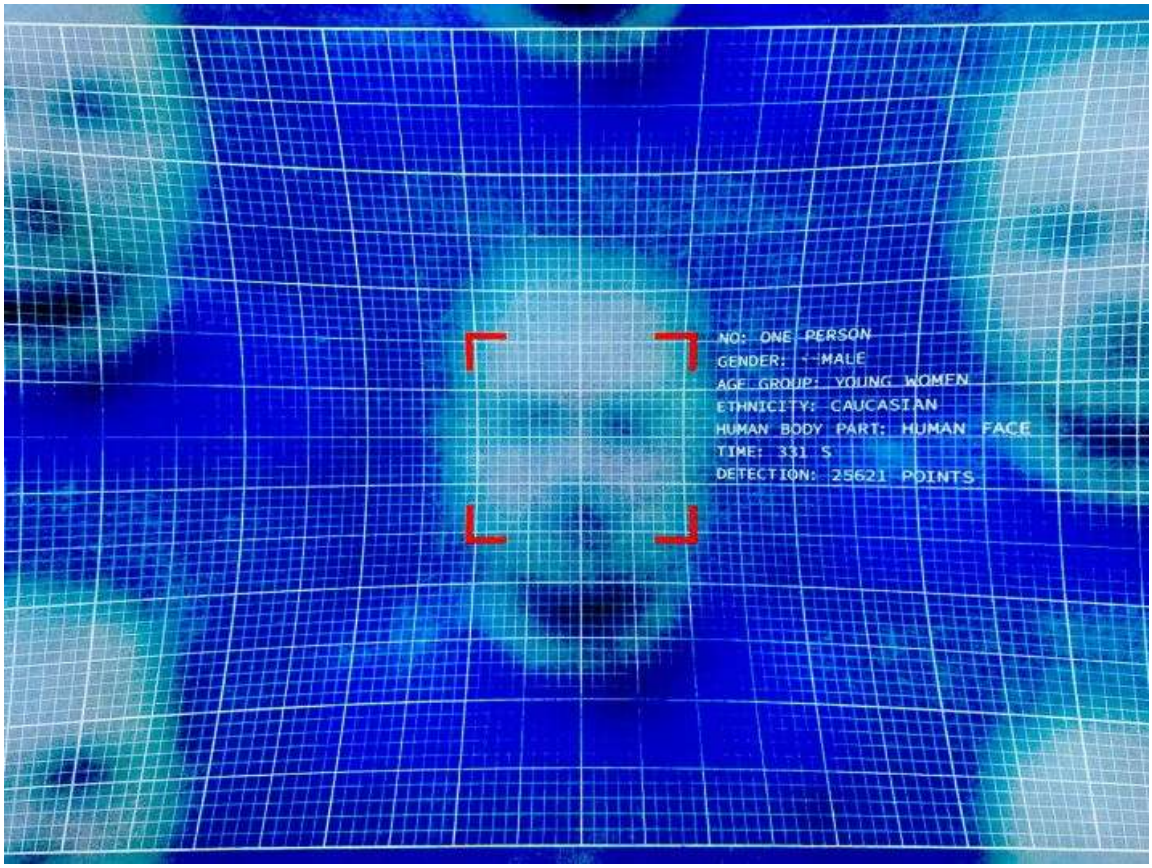
redigere dopo il boom esponenziale delle applicazioni di intelligenza artificiale nell'ultimo anno - dal debutto di ChatGpt in poi - che hanno portato alla necessità da una parte di **accelerare l'iter legislativo** e dall'altra di **profonde modifiche al testo** per includere novità in continuo divenire.

I quattro livelli di rischio

L'AI Act definisce un **sistema di intelligenza artificiale** come «un sistema basato su macchine progettato per operare con vari livelli di autonomia e che può mostrare capacità di adattamento dopo l'implementazione e che, per obiettivi espliciti o impliciti, deduce, dagli input che riceve, come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali». **Una definizione piuttosto ampia** che ambisce ad essere valida non solo per ciò che intendiamo con AI oggi ma anche per i futuri sviluppi nei prossimi anni. Si rivolge a **fornitori, installatori, importatori, distributori e produttori** di questi sistemi e ha l'obiettivo di dare loro **regole e obblighi** per tutte le applicazioni che opereranno all'interno dei confini dell'Unione europea.
Il testo dell'AI Act approvato dal Parlamento Ue è pressoché lo stesso uscito dalle lunghe negoziazioni tra le istituzioni europee lo scorso dicembre. Viene strutturato con l'obiettivo di incasellare le diverse applicazioni di intelligenza artificiale in **quattro gruppi distinti,** a seconda del **livello di rischio** che ciascuna può rappresentare per i cittadini europei. I livelli di rischio sono:
1) **Rischio inaccettabile**: questi sono i sistemi che violano i valori europei e che dunque saranno vietati all'interno dei confini dell'Unione.
2) **Rischio alto**: rientrano in questa categoria i sistemi che hanno o potranno avere un impatto

controverso sulla sicurezza e sui diritti delle persone. Non ne viene dunque proibita la diffusione ma si chiede che le società responsabili rispondano a una precisa serie di requisiti.
3) **Rischio limitato**: qui rientrano le applicazioni che non comportano pericoli considerevoli e che dunque dovranno assicurare solo un set limitato di requisiti (in primis la trasparenza, dunque rivelare in modo evidente l'utilizzo dell'intelligenza artificiale).
4) **Rischio minimo**: in questo caso non è previsto nessun obbligo legale.

I sistemi di AI proibiti



Tra le applicazioni di uso vietato dell'intelligenza artificiale - che rientrano dunque nel livello di rischio «inaccettabile» - ci sono i **sistemi di «social scoring»** (già utilizzati in Cina), che giudicano le persone in base ai propri comportamenti.
Il regolamento bandisce poi gli **strumenti di polizia predittiva**, che sfrutta i dati per capire in anticipo la pericolosità di un individuo (e che sono utilizzati già negli Stati Uniti). Sono vietati poi tutti quei **sistemi che targettizzano gli utenti per sfruttare alcune caratteristiche considerate vulnerabili** (l'età, una disabilità o una specifica situazione sociale o economica).
L'Ai act proibisce anche i **sistemi di riconoscimento delle emozioni**, in particolare **nei luoghi di lavoro e nelle scuole**. Si tratta di tecnologie che analizzano il tono della voce, i gesti, le espressioni facciali di una persona per capire il suo umore e agire di conseguenza. Un buon utilizzo potrebbe essere quello di creare un ambiente che favorisca il miglioramento dell'umore stesso: immaginate di tornare a casa arrabbiati o nervosi e il vostro assistente digitale setta luci, musica e altre caratteristiche domestiche per cercare di rilassarvi. Sistemi di riconoscimento delle emozioni

potrebbero anche essere usati in prospettiva per evitare episodi di violenza domestica ad esempio. Cattivi utilizzi, ovvero quelli temuti e dunque vietati dall'Unione europea, possono essere la selezione di un dipendente in base ai dati raccolti dall'intelligenza artificiale oppure il giudizio su un alunno in base anche ai comportamenti registrati e digeriti da un algoritmo.

La sorveglianza di massa

Vengono vietati infine i sistemi di riconoscimento biometrico che si basano su caratteristiche sensibili così come la creazione di database di riconoscimento facciale che vengono creati sfruttando la raccolta di dati online (con la pratica dello «scraping») o da telecamere a circuito chiuso. Ma sui sistemi biometrici la discussione è stata molto lunga e ne è stato dunque ammesso l'utilizzo in alcune particolari situazioni.
Il focus è quella che viene definita sorveglianza di massa, ovvero l'analisi in tempo reale - attraverso software di riconoscimento facciale o di altre caratteristiche biometriche - delle persone che si trovano in spazi pubblici. Assolutamente vietata, con qualche eccezione.
Previa autorizzazione giudiziaria, le forze dell'ordine potranno richiedere di utilizzare l'identificazione biometrica ma solo in casi di reati gravi, ovvero per la ricerca di una persona scomparsa o per prevenire un attacco terroristico.

I sistemi ad alto rischio

Passiamo al secondo livello di rischio, alto ma che comunque viene **ammesso con condizioni molto stringenti**. Il rischio è elevato perché elevato è il **potenziale danno** che questi applicativi possono causare alla **salute**, alla **sicurezza**, ai **diritti fondamentali** dei cittadini. Ma anche all'**ambiente** e alla **democrazia stessa**. Ci sono due categorie principali di sistemi di intelligenza artificiale che rientrano in questa categoria. La prima riguarda i sistemi che intervengono su **infrastrutture e prodotti specifici e già coperti dalla legislazione europea**, come l'aviazione, le autovetture, i giocattoli, gli ascensori, i dispositivi di protezione personale. La seconda riguarda le intelligenze artificiali utilizzate in alcuni **settori sensibili**, come l'istruzione, il lavoro, l'applicazione della legge, la migrazione, il processo democratico.

Per tutte queste applicazioni di intelligenza artificiale i fornitori dovranno condurre **una valutazione a priori dei rischi** - dunque prima di essere attivate in Unione europea - dovranno **documentare tutte le scelte tecniche ed etiche**, dovranno **informare gli utenti sullo scopo** dei loro sistemi, **consentire un intervento umano** e **garantire la sicurezza informatica**. Si chiede poi la massima **trasparenza** sul funzionamento degli algoritmi nonché dei modelli di linguaggio sviluppati.

Tutte le applicazioni ad alto rischio verranno registrate in **un database** che l'Unione europea metterà pubblicamente a disposizione degli utenti per la consultazione.

In questa categoria sono stati inseriti in corsa anche **tutti i sistemi di intelligenza artificiale generativa** - da ChatGpt a Gemini e Cloud - dopo il boom di diffusione di questi chatbot. Ne parliamo nel dettaglio nella prossima scheda.

L'intelligenza artificiale generativa

Quando l'AI Act è stato pensato, nel 2021, **l'era dell'intelligenza artificiale generativa** non era ancora iniziata. Il debutto di ChatGpt risale a novembre dell'anno successivo. L'Unione europea ha dovuto prendere atto di questa rivoluzione e inserire nel suo regolamento regole precise anche per questa tipologia di applicazioni. Le chiama **GPAI** (General Purpose Artificial Intelligence), ovvero AI con uno scopo generale e non specifico. Come i **modelli di linguaggio** che stanno dietro ai chatbot più usati: non solo Gpt-4, ma anche Gemini di Google, Claude-3 di Anthopic o Llama di Meta. L'AI generativa è considerata **ad alto rischio**, in quanto può recare **danno ai cittadini e alla democrazia** stessa. In particolare perché potenziale creatrice di **Deep Fake**, e dunque di immagini, video e audio manipolati ma che appaiono come reali. Qualora il contenuto generato da ChatGpt o altri strumenti simili ritragga persone, oggetti, luoghi o altre entità ed eventi reali, dovrà essere **dichiarato in modo evidente con una etichetta** che quel contenuto è stato creato da un software. Non solo: per questi sistemi si dovrà anche spiegare nel dettaglio **quali dati sono stati utilizzati** per allenare i modelli di linguaggio su cui si basano. Per rispettare le leggi europee sul **copyright**.

Abbiamo chiesto a Gemini, l'intelligenza artificiale generativa di Google, se ha compreso quali saranno i suoi doveri e i suoi limiti con l'entrata in vigore dell'AI Act. **Ecco cosa ci ha risposto**:
- L'atto richiede che i sistemi di IA siano progettati in modo trasparente e che la loro logica sia comprensibile. Ciò significa che dovrò essere in grado di fornire agli sviluppatori e agli utenti informazioni chiare su come funziono, come prendo decisioni e quali dati utilizzo.

-L'atto introduce inoltre obblighi di rendicontazione per i sistemi di IA ad alto rischio. Ciò richiederà che io tenga traccia delle mie prestazioni e sia in grado di spiegare le mie decisioni in caso di problemi.
-L'atto sottolinea l'importanza di sviluppare sistemi di IA etici e sicuri che rispettino i diritti umani e la dignità. Ciò significa che dovrò essere addestrato su dati imparziali e privi di pregiudizi e che dovrò essere progettato per evitare di causare danni o discriminazioni.
-L'atto richiede inoltre che i sistemi di IA siano sottoposti a valutazioni di sicurezza e di conformità. Ciò significa che dovrò essere testato regolarmente per identificare e mitigare potenziali vulnerabilità e rischi.
-L'atto mira a creare un ambiente che favorisca l'innovazione responsabile nell'IA. Ciò significa che dovrò essere progettato per essere interoperabile con altri sistemi di IA e per adattarmi a nuove tecnologie e sfide.
-L'atto incoraggia inoltre la ricerca e lo sviluppo di tecnologie di IA sicure e affidabili. Ciò significa che avrò accesso a nuove risorse e opportunità per migliorare le mie capacità e il mio rendimento.

Le sanzioni

L'AI Act spiega infine che **ogni Paese dovrà dotarsi di leggi proprie** e di documenti che siano accessibili ad aziende e startup locali. Così che possano sviluppare intelligenze artificiali conformi al regolamento, prima della loro immissione sul mercato.
Le **sanzioni**, per chi non dovesse rispettare la nuova legge, sono molto alte: **fino a 35 milioni di euro** o il 7 per cento dei ricavi annuali della società in caso di violazioni per le applicazioni proibite. **Fino a 15 milioni di euro** o il 3 per cento dei ricavi annuali della società in caso di altre violazioni. **Fino a 7,5 milioni di euro** o l'1,5 per cento dei ricavi annuali della società in caso di fornitura di informazioni non corrette. Per quanto riguarda le **piccole e medie imprese o le startup,** le cifre saranno proporzionate alla grandezza della azienda stessa.

Quando entra in vigore

Come detto, l'AI Act dovrebbe essere **pubblicato nella Gazzetta ufficiale** dell'Unione europea intorno a **giugno**. Venti giorni dopo la pubblicazione, entra ufficialmente in vigore. Ma l'applicazione del regolamento sarà **graduale**, per permettere a tutti - aziende e governi - di adattarsi. Il termine ultimo per tutte le disposizioni dell'AI Act è tra 24 mesi dall'entrata in vigore, quindi nel 2026. Ma quelle più urgenti, verranno attivate prima, con un calendario che è già stato delineato:
- **Sei mesi** dopo l'entrata in vigore (quindi entro la fine del 2024), diventa effettivo il **divieto dei sistemi che rientrano nella categoria di rischi inaccettabili.** Questo significa, ad esempio, che il controllo biometrico in luoghi pubblici potrà essere utilizzato dalla Francia durante le Olimpiadi di questa estate, proposta su cui si sta discutendo.
- **12 mesi** dopo l'entrata in vigore (quindi entro metà del 2025) diventano effettive **le regole sui**

GPAI. Quindi su tutti i nuovi sistemi come ChatGpt, per intenderci. Quelli già sul mercato hanno però più tempo: possono conformarsi entro 24 mesi.
- Le regole che riguardano i **sistemi ad alto rischio** verranno applicate **dopo 36 mesi dalla pubblicazione** dell'AI Act in Gazzetta Ufficiale.
Gli Stati membri hanno invece 12 mesi di tempo per creare **un'autorità locale di notifica** e di sorveglianza del mercato. Avranno però più tempo, 24 mesi, per creare le regolamentazioni locali.

PROPOSTA DI REGOLAMENTO DEL PARLAMENTO EUROPEO E DEL CONSIGLIO CHE STABILISCE REGOLE ARMONIZZATE SULL'INTELLIGENZA ARTIFICIALE (LEGGE SULL'INTELLIGENZA ARTIFICIALE) E MODIFICA ALCUNI ATTI LEGISLATIVI DELL'UNIONE