Università degli Studi di Ferrara
Agenda 17
Webmagazine del Laboratorio Design of Science

DOSSIER AI ACT L'impatto della normativa europea sull'intelligenza artificiale: una garanzia di sicurezza o un ostacolo all'innovazione?

di Enrico Maestri, Professore associato di Filosofia del diritto Università di Ferrara, Titolare degli insegnamenti di Diritto informatico, di Metodologia e Logica giuridica, di Etica e diritto dell'intelligenza artificiale

1. Introduzione

Il Consiglio dell'UE ha approvato unanimemente il 21 maggio 2024 l'AI Act, il primo regolamento compiuto al mondo sull'intelligenza artificiale

. Questo atto storico stabilisce norme armonizzate per la progettazione, lo sviluppo e l'utilizzo dei sistemi di intelligenza artificiale nell'Unione Europea, basandosi su un approccio proporzionato al rischio.

Proposto originariamente dalla Commissione Europea nell'aprile 2021 per affrontare le sfide poste dai rapidi progressi tecnologici e dai potenziali rischi associati all'intelligenza artificiale, il processo legislativo ha subito una modifica nel dicembre 2022 con l'introduzione di ChatGPT. Ciò ha reso necessario includere norme specifiche per l'intelligenza artificiale generativa. La commissione per il mercato interno e la protezione dei consumatori del Parlamento europeo ha adottato la legge il 13 marzo 2024, e sottoposta al voto in plenaria previsto per il 10 e 11 aprile 2024.

L'AI Act è il primo provvedimento normativo al mondo sull'IA che contiene un insieme organico di regole per tutelare i diritti della persona, con un approccio umano-centrico e basato sul rischio.

L'intelligenza artificiale ha il potenziale per trasformare diversi ambiti, penetrando in tutte le industrie e i settori.
L'AI Act rappresenta un'iniziativa storica volta a rendere l'UE un leader globale nel campo dell'IA etica e centrata sull'uomo, promuovendo contemporaneamente l'innovazione e la competitività. La legge sull'IA fornisce indicazioni all'interno del suo ampio quadro, ma permane l'incertezza sull'attuazione pratica della stessa. Alcune questioni riguardano l'applicazione delle norme ai

Pubblicato su: https://www.agenda17.it/

modelli di intelligenza artificiale generativa, come ChatGPT, che sollevano nuove questioni relative al diritto d'autore e alla proprietà intellettuale. La Commissione Europea ha suggerito una revisione della Direttiva UE sul diritto d'autore per tener conto dei progressi tecnologici dell'intelligenza artificiale generativa e del loro impatto sulla proprietà intellettuale e industriale.

Negli ultimi mesi, il dibattito sulla ricerca e sviluppo dell'intelligenza artificiale (IA) e le sue implicazioni socioeconomiche è diventato sempre più acceso.
L'ascesa dell'IA generativa, con i suoi chatbot come ChatGPT di OpenAI e Gemini di Google, e i modelli linguistici di grandi dimensioni (LLM) sottostanti (come LaMDA di Google, LLaMA di Meta, GPT-4 di OpenAI e PaLM di Google), ha sollevato numerose preoccupazioni.

Critici e studiosi lanciano l'allarme sui potenziali rischi associati all'intelligenza artificiale (IA), in particolare ai Large Language Models (LLM). La facilità con cui questi modelli possono generare testi realistici desta preoccupazioni in merito alla diffusione di disinformazione e contenuti ingannevoli, alla manipolazione delle opinioni online e alla perdita di fiducia nelle informazioni digitali.

L'impatto dell'IA sul mercato del lavoro è un'altra fonte di preoccupazione. L'automazione spinta da questa tecnologia potrebbe portare a una significativa perdita di posti di lavoro in svariati settori, con conseguenze sociali ed economiche potenzialmente gravi.
Figure di spicco del mondo scientifico e tecnologico hanno espresso forti allarmi in merito a questi rischi. Sir Patrick Vallance, ex capo scienziato del Regno Unito, paragona la rivoluzione dell'IA a quella industriale in termini di impatto sul mercato del lavoro. IBM ha annunciato la sospensione delle assunzioni per ruoli potenzialmente automatizzabili dall'IA.
Il "Future of Jobs Report" del 2023 del World Economic Forum evidenzia il rischio di una massiccia distruzione di posti di lavoro a causa dell'automazione. Geoffrey Hinton, pioniere dell'IA, ha lasciato Google per via dei pericoli associati a questa tecnologia. Elon Musk e altri leader del settore chiedono una moratoria sullo sviluppo dell'IA per valutare e mitigare i rischi.
Il Center for AI Safety definisce l'IA un "grave rischio di estinzione" per l'umanità, sottolineando la necessità di una riflessione profonda e di un'azione urgente per garantire uno sviluppo responsabile e sicuro di questa tecnologia.

La discussione sui rischi dell'IA è complessa e sfaccettata, e richiede un approccio equilibrato che consideri i potenziali benefici di questa tecnologia insieme ai rischi e alle sfide che essa pone.

Pubblicato su: https://www.agenda17.it/

Queste preoccupazioni hanno spinto a richiedere una regolamentazione globale e internazionale dell'IA. Diverse istituzioni e governi in tutto il mondo hanno già iniziato ad adottare misure in questo senso. La prima mossa definitiva in tal senso è stata presa dal Consiglio dell'UE che ha dato il via libera all'AI Act, il primo regolamento completo al mondo sull'intelligenza artificiale.

2. Struttura e contenuti dell'AI Act

L'AI Act, frutto di un'ampia consultazione con le parti interessate, incorpora i contributi del Gruppo di esperti di alto livello sull'intelligenza artificiale (vedi Linee guida etiche per un'intelligenza artificiale affidabile del 2019) e si basa sul Libro bianco dell'UE (Commissione europea, 2020).

Inoltre, l'AI Act integra il Regolamento generale sulla protezione dei dati (GDPR) ed è in linea con il Digital Services Act (DSA), il Digital Markets Act (DMA) e altre iniziative normative, come il regolamento sui prodotti meccanici, il Data Governance Act, la direttiva sulla responsabilità dell'intelligenza artificiale, la direttiva rivista sulla responsabilità del prodotto e vari quadri settoriali sulla sicurezza dei prodotti.

L'AI Act è composto da 12 titoli, 85 articoli e nove allegati. Il regolamento direttamente applicabile negli Stati membri è impostato su una architettura di rischi**.** Maggiore è il rischio, maggiori sono le responsabilità e i paletti per chi sviluppa o adopera sistemi di intelligenza artificiale. Fino alle applicazioni considerate troppo pericolose per essere autorizzate. Gli unici casi che non ricadono sotto l'ombrello dell'AI Act sono le tecnologie adoperate per scopi militari e per quelli di ricerca.

L'articolo 1 stabilisce norme armonizzate per l'immissione sul mercato, la messa in servizio e l'uso dei sistemi di IA nell'Unione europea. Sono previsti divieti per pratiche specifiche, requisiti per sistemi ad alto rischio e norme sulla trasparenza. Regole specifiche riguardano sistemi destinati a interagire con persone, sistemi di riconoscimento emotivo e sistemi di categorizzazione biometrica. Sono inoltre incluse regole sul monitoraggio e la sorveglianza del mercato.

Il regolamento si applica ai fornitori che mettono sul mercato o mettono in servizio sistemi di IA nell'Unione, indipendentemente dalla loro ubicazione (articolo 2).
Ciò significa che l'AI Act si applica ad attori pubblici e privati, all'interno e all'esterno dell'UE, quando i sistemi di intelligenza artificiale hanno un impatto sui cittadini dell'UE.
Secondo l'articolo 3 e l'Allegato III, i sistemi di intelligenza artificiale comprendono approcci di machine learning, di deep learning, di logica e rappresentazione della conoscenza e approcci statistici.

Da notare che nel 2023, di fronte al dirompente e allo strabiliante sviluppo dei sistemi di machine learning e di deep learning, il Parlamento europeo ha modificato l'elenco dei sistemi di IA intrusivi e discriminatori e quindi vietati per includere anche i sistemi di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico, i sistemi di identificazione biometrica remota "post" [...], i sistemi di categorizzazione biometrica che utilizzano caratteristiche sensibili (ad esempio, genere, razza, etnia, stato di cittadinanza, religione, orientamento politico), i sistemi di polizia predittiva [...], i sistemi di riconoscimento delle emozioni nelle forze dell'ordine, nella gestione delle frontiere, sul posto di lavoro e nelle istituzioni educative e la rimozione indiscriminata di dati biometrici dai social media o da filmati CCTV per creare database di riconoscimento facciale.

Si noti che la maggior parte dei sistemi sopra elencati erano precedentemente etichettati come ad alto rischio e quindi consentiti; la polizia predittiva, tuttavia, non era ancora inclusa nella bozza dell'AIA, né nella categoria inaccettabile né in quella ad alto rischio.

Inoltre, il Parlamento ha inteso ampliare la classificazione delle aree ad alto rischio per includere "danni alla salute delle persone, alla sicurezza, ai diritti fondamentali o all'ambiente".

Ha inoltre inteso aggiungere i sistemi utilizzati per influenzare o manipolare gli elettori nelle campagne politiche (ad esempio, il micro-targeting politico) e i sistemi di raccomandazione utilizzati dalle piattaforme di social networking all'elenco dei sistemi di intelligenza artificiale ad alto rischio.

Inoltre, il Parlamento ha voluto includere disposizioni aggiuntive per i fornitori di modelli di fondazione generativi, come gli LLM/GPT. Tali entità avrebbero dovuto rispettare una maggiore trasparenza e altri requisiti legali, ad esempio rivelando agli utenti che il contenuto è stato generato da detti sistemi di intelligenza artificiale, progettando modelli in modo da impedire la generazione di contenuti illegali e fornendo una panoramica dei dati protetti da copyright utilizzati per la formazione.

Infine, il Parlamento ha richiesto di rafforzare il diritto dei cittadini di presentare reclami sui sistemi di intelligenza artificiale e di ricevere spiegazioni sulle decisioni basate su sistemi di intelligenza artificiale ad alto rischio che incidono in modo significativo sui loro diritti.

Pubblicato su: https://www.agenda17.it/

L'AI Act si basa su alcuni principi chiave che ne guidano l'applicazione e la governance:

- Intervento e sorveglianza umani (articolo 14): l'uomo mantiene un ruolo centrale nel controllo e nella supervisione dei sistemi di intelligenza artificiale. Questo principio mira a garantire che l'uomo sia sempre in grado di comprendere e gestire i sistemi di IA, prevenendo potenziali rischi o abusi.
- Robustezza tecnica e sicurezza (articolo 15): i sistemi di IA devono essere progettati e sviluppati in modo robusto e sicuro, adottando misure adeguate per prevenire malfunzionamenti, attacchi informatici e altri rischi per la sicurezza.
- Vita privata e governance dei dati (articolo 10): la protezione dei dati personali è un principio fondamentale dell'AI Act. I sistemi di IA devono trattare i dati personali in modo conforme al Regolamento generale sulla protezione dei dati (GDPR) e ad altri principi di privacy.
- Trasparenza (articoli 13 e 50): gli utenti devono avere accesso a informazioni chiare e comprensibili su come funzionano i sistemi di IA e su come vengono utilizzati i loro dati. Questo principio mira a promuovere la fiducia e la responsabilità nell'utilizzo dell'IA.
- Diversità, non discriminazione ed equità (articolo 27): i sistemi di IA devono essere progettati e utilizzati in modo da evitare discriminazioni e promuovere l'equità. Questo principio mira a garantire che i sistemi di IA non rafforzino o perpetuino pregiudizi esistenti nella società.
- Benessere sociale e ambientale (articolo 27): lo sviluppo e l'utilizzo dei sistemi di IA devono essere orientati al benessere sociale e ambientale. Questo principio mira a garantire che l'IA sia utilizzata in modo da apportare benefici alla società e all'ambiente, non causando danni.
- Responsabilità dei fornitori di sistemi IA (articolo 25): i fornitori di sistemi di IA sono responsabili della conformità ai requisiti dell'AI Act. Questo principio mira a garantire che le aziende che sviluppano e commercializzano sistemi di IA siano responsabili delle loro azioni e dei potenziali rischi associati ai loro prodotti.

Questi principi chiave costituiscono la base dell'AI Act e hanno l'obiettivo di promuovere uno sviluppo e un utilizzo dell'intelligenza artificiale sicuro, etico e responsabile.

Pubblicato su: https://www.agenda17.it/

3. Aspetti essenziali della legge sull'intelligenza artificiale (AI Act)

3.1.: una definizione ampia dei sistemi di intelligenza artificiale

La legge sull'intelligenza artificiale definisce un sistema di intelligenza artificiale come "un sistema basato su macchine e progettato per operare con diversi livelli di autonomia e che può mostrare adattività dopo l'implementazione, deducendo, dagli input ricevuti, come generare output come previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali" (considerando 6, art. 3), con ciò seguendo l'ultima definizione elaborata dall'OCSE.

Le parole chiave in questa definizione sono "deduzione" e "autonomia", che distinguono chiaramente un sistema di intelligenza artificiale da qualsiasi altro software in cui l'output è predeterminato (se x allora y) da un algoritmo rigoroso (sistema esperto). Questa definizione è appositamente ampia per garantire che la legge sull'IA non diventi obsoleta in futuro. Si discosta nettamente dalla definizione originale dei sistemi esperti di intelligenza artificiale, che legava il concetto a un elenco predefinito di tecnologie e metodi, adottando un approccio tecnologicamente neutro e uniforme.

Parafrasata in termini epistemologici, la definizione ampia adottata dal legislatore europeo è un sintomo della consapevolezza che se ammettiamo che l'intelligenza artificiale emuli (copi, simuli) il pensiero umano, dobbiamo tener conto che quest'ultimo è solitamente impreciso, difficilmente certo e più probabilmente plausibile, ragionevole forse, ma quasi mai razionale; è un pensiero compromissorio e defettibile: sostanzialmente non logico (che non significa illogico). Quindi, se l'intelligenza artificiale emula questo pensiero, essa genera esattamente gli stessi problemi che sono soliti degli uomini e non delle macchine.

Pertanto, l'intelligenza artificiale che emula il pensiero umano può generare problemi simili a quelli che gli esseri umani potrebbero affrontare.

Oggi però siamo già oltre l'emulazione: l'intelligenza artificiale non è più (soltanto) l'emulazione del pensiero umano, ma offre molte tecniche per l'elaborazione di pensieri computazionali basati su premesse formalmente certe desunte dalla misurazione dei fenomeni del mondo.

Queste tecniche si basano sull'applicazione di principi logici, matematici e sulle leggi del calcolo delle probabilità e consentono di risolvere problemi in cui l'intervento "pensante" della macchina è capace di andare molto oltre le capacità dell'uomo. L'uomo lascia dunque alla macchina un

patrimonio di modelli e di dati, chiedendo alla macchina di fare il passo in più che egli non può compiere a causa dei suoi limiti fisici e cognitivi.

Il salto di qualità è costituito dall'algoritmo d'autoapprendimento (machine learning e deep learning), che scoperchia il vaso di Pandora, che affranca il processo algoritmico dai vincoli di predeterminazione impressi dal programmatore e che lo rende capace di percorso autonomo verso esiti e scelte "autonome" e non predicabili. E se è vero che la A.I. è una macchina che si nutre di dati, non è sempre vero che essa resta inevitabilmente condizionata dai dati che si sceglie di darle in pasto. Si dice, del resto, che basterebbe un algoritmo con autoapprendimento ospitato da un computer con accesso ad internet perché quella A.I. possa conquistare il mondo (comandando delle entità robotiche o corrompendo degli umani).

Una A.I. self-learning sufficientemente avanzata, dunque, può sviluppare la capacità di accedere a dataset anche diversi da quelli immaginati dal suo creatore.

Sebbene molti gruppi di stakeholder e studiosi in generale sostengano l'AI Act, chiedono anche diverse modifiche, tra cui, ma non solo, il perfezionamento della definizione di sistemi di IA, l'ampliamento dell'elenco dei sistemi di IA vietati e ad alto rischio, una migliore protezione dei diritti fondamentali, il rafforzamento dei sistemi di governance e meccanismi di ricorso e garantire un adeguato controllo democratico e giudiziario dell'attuazione dell'AI Act.

Organizzazioni imprenditoriali criticano il fatto che la definizione di sistemi di intelligenza artificiale dell'AI Act è troppo ampia, che non è chiaro il trattamento dei componenti dei sistemi di IA più grandi e che l'AI Act potrebbe portare a un'eccessiva regolamentazione. Sono preoccupate per i costi amministrativi delle procedure introdotte, come l'implementazione di sistemi di gestione del rischio, pratiche di gestione dei dati, documentazione tecnica e altri obblighi.

Le organizzazioni per i diritti civili chiedono un divieto di utilizzo arbitrario dei dati biometrici e restrizioni più ampie sull'uso dei sistemi di IA, mentre le organizzazioni per i diritti civili chiedono valutazioni più approfondite dei rischi e degli impatti sui diritti umani e sulla trasparenza.

Le associazioni di difesa dei consumatori sostengono che l'AI Act dovrebbe avere una portata più ampia e imporre obblighi essenziali di trasparenza, responsabilità ed equità a tutti i fornitori di sistemi di intelligenza artificiale, oltre a vietare in modo più completo pratiche dannose e pericolose, come i sistemi di identificazione biometrica remota negli spazi pubblici e privati e l'utilizzo di social scoring da parte degli enti. Esortano inoltre la Commissione a garantire ai consumatori diritti individuali, mezzi di ricorso efficaci e meccanismi di ricorso.

Pubblicato su: https://www.agenda17.it/

3.2. Carattere giuridicamente vincolante

L'AI Act è la prima iniziativa governativa significativa al mondo volta ad affrontare – e idealmente mitigare – gli impatti potenzialmente negativi dei sistemi di intelligenza artificiale. Pertanto, molti lo vedono come una pietra miliare nella regolamentazione di tali tecnologie informatiche.

Ciò che distingue l'AI Act da altre iniziative normative è il suo carattere giuridicamente vincolante, ovvero il carattere di legge dura e l'inclusione di requisiti obbligatori per sviluppatori, operatori e utenti di IA. Segna quindi un significativo allontanamento dalle precedenti iniziative di autoregolamentazione, vale a dire iniziative di soft law in materia di etica dell'IA in Europa e altrove.

Molti ricercatori evidenziano le caratteristiche problematiche dei sistemi di intelligenza artificiale, tra cui complessità, interconnettività, attenzione alla correlazione rispetto alla causalità, capacità di apprendimento e adattamento continuo, comportamento (semi-)autonomo e opacità, nota anche come "black box". Tali caratteristiche hanno portato a questioni etiche associate a tali sistemi, come la sostituzione degli esseri umani con le macchine, interfacce cervello-computer e miglioramento umano, problemi di sicurezza, responsabilità, privacy e protezione dei dati, manipolazione e discriminazione algoritmica, come profilazione, micro-targeting, nudging politico e diffusione di disinformazione.

Considerando le implicazioni sociali dell'intelligenza artificiale, la regolamentazione di tali tecnologie sembra giustificata, ma molti paesi nel mondo stanno affrontando un vuoto normativo. L'obiettivo principale dell'AI Act è colmare questa lacuna e tenere il passo con gli ultimi progressi tecnologici.

È interessante notare che l'UE ha inizialmente seguito un approccio di soft law con i suoi orientamenti etici non vincolanti per un'intelligenza artificiale affidabile e raccomandazioni su politiche e investimenti (gruppo di esperti di alto livello sull'intelligenza artificiale, 2019), ma ha poi spostato la rotta verso un approccio legislativo di hard law: un approccio quindi che ha chiesto l'adozione di un nuovo quadro normativo sull'IA.

Queste norme di nuova creazione, giuridicamente vincolanti, integrano e seguono la logica delle norme UE esistenti sulla sicurezza dei prodotti e sulle norme di protezione.

Tuttavia, la strategia dell'UE è nettamente diversa dall'approccio adottato dagli Stati Uniti (e da altri paesi), che segue un approccio laissez-faire e di non intervento nei confronti della regolamentazione dell'IA (eccezioni degne di nota sono l'ordine esecutivo emesso da Biden su "Safe, Secure e un'intelligenza artificiale affidabile" e l'Algorithmic Accountability Act, attualmente dibattuto al Congresso, ma dato il clima politico esistente negli Stati Uniti, l'atto probabilmente non passerà.

È lodevole anche il fatto che l'AI Act presuppone un ordine mondiale post-westfaliano e che sia extraterritoriale. Cioè, non importa dove si trova un fornitore o operatore di un sistema di IA purché i suoi servizi abbiano un impatto sui cittadini dell'UE (art. 2 dell'AI Act).
Di conseguenza, le lacune geografiche non possono essere sfruttate per eludere la portata dell'AI Act - ad esempio, da parte di un'azienda che si trasferisce in un paese con una regolamentazione meno rigorosa - garantendo così la protezione tra gli Stati membri.

In quanto tale, l'AI Act può potenzialmente estendere il cosiddetto "effetto Bruxelles" ad altri ambiti politici. L'effetto presuppone che le aziende rispettino le leggi e i regolamenti dell'UE, come il GDPR, anche negli Stati non membri dell'UE perché è più pratico avere un approccio unico a livello globale, consentendo all'UE di estendere de facto - anche se non de jure - la sua portata norme a livello internazionale attraverso meccanismi di regolamentazione del mercato.

La speranza è che la "leadership basata sull'esempio" dell'UE possa innescare una legislazione simile sull'IA in tutto il mondo, ad esempio, la legislazione statale statunitense come la CCPA/CPRA (nell'area delle leggi sulla privacy e sulla protezione dei dati).

Successivamente, l'obiettivo dell'AI Act è garantire che lo sviluppo dell'IA nell'UE sia "eticamente corretto, legalmente accettabile, socialmente equo e sostenibile dal punto di vista ambientale, con una visione dell'IA che cerchi di sostenere [cioè servire] l'economia, la società, e l'ambiente". È quindi in linea con la visione (business case) secondo cui l'etica avvantaggia i mercati, e non viceversa.

Sebbene questo sia un aspetto cruciale, gli studiosi sottolineano anche che l'AI Act potrebbe svolgere un lavoro migliore proteggendo i diritti dei consumatori e altri diritti fondamentali e fornendo misure per rimediare a possibili danni o perdite causati dai sistemi di intelligenza artificiale.

Pubblicato su: https://www.agenda17.it/

3.3.Classificazione dei sistemi di IA

II Regolamento definisce tre categorie di rischio per i sistemi di IA: inaccettabile, alto e basso. 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.

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.

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).

Pubblicato su: https://www.agenda17.it/

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 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.

Pubblicato su: https://www.agenda17.it/

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.

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.

Pubblicato su: https://www.agenda17.it/

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.

Pubblicato su: https://www.agenda17.it/

3.4. Elenco chiuso di sistemi AI vietati e meccanismo sfumato per l'identificazione biometrica remota in tempo reale

Come abbiamo visto, la legge sull'IA presenta un elenco chiuso di pratiche interdette di intelligenza artificiale, tra cui: l'utilizzo di tecniche subliminali o manipolative per distorcere il comportamento; lo sfruttamento delle vulnerabilità di individui o gruppi specifici; sistemi di categorizzazione biometrica individuale tranne per l'etichettatura legale dei dati biometrici acquisiti dalle forze dell'ordine; sistemi di punteggio sociale; identificazione biometrica remota in tempo reale in luoghi pubblici a fini di contrasto; attività di polizia predittiva basate solo sulla profilazione tranne se supportate da valutazioni umane basate su fatti oggettivi legati alla criminalità; database di riconoscimento facciale non mirato; inferenza delle emozioni sul luogo di lavoro o in istituti scolastici, tranne per motivi medici o di sicurezza.

Il divieto dell'identificazione biometrica in tempo reale per fini di contrasto ha generato vivaci discussioni nelle istituzioni europee. Tuttavia, l'eccezione si applica a specifici scopi come la ricerca di vittime di tratta umana o la lotta contro il terrorismo, richiedendo valutazioni rigorose, misure tecniche e organizzative, notifiche adeguate e una giustificazione giuridica.

I critici elogiano anche la capacità teorica dell'AI Act di affrontare le tre principali categorie di rischio relative ai sistemi di intelligenza artificiale: rischi relativi alla qualità dei dati (correttezza, tempestività e rappresentatività dei dati), rischi di discriminazione (discriminazione algoritmica, ad esempio dovuta a distorsioni di campionamento come la sotto-rappresentanza di alcuni gruppi) e rischi di innovazione (rischio di eccessiva regolamentazione e rischio di blocco, ad esempio a causa dei diritti di proprietà intellettuale).

L'AI Act si concentra principalmente sulla qualità dei dati (rischi) e ha il potenziale per istituire un regime di qualità indipendente per i dati di addestramento. Richiede, tra l'altro, che i dati siano "pertinenti, rappresentativi, privi di errori e completi"; richiede inoltre un equilibrio dei set di dati tra i membri dei gruppi protetti.
L'AI Act rappresenta quindi un passo essenziale nella giusta direzione, ma, come mostrato di seguito, potrebbe fare molto di più, soprattutto per quanto riguarda i rischi di discriminazione, cioè nell'affrontare i pregiudizi algoritmici e la discriminazione.

Pubblicato su: https://www.agenda17.it/

3.5.Regole ed eccezioni alla qualificazione dei sistemi di IA ad alto rischio

La legge sull' AI introduce un'eccezione significativa a tale classificazione: i sistemi di IA della seconda categoria ad alto rischio (allegato III) che non presentano un rischio significativo per la salute, la sicurezza o i diritti fondamentali delle persone fisiche, non saranno considerati ad alto rischio.
Questo si applica se il sistema svolge compiti specifici, migliora risultati umani preesistenti, rileva modelli decisionali o deviazioni da modelli precedenti senza sostituire o influenzare la valutazione umana senza una revisione adeguata, o se svolge attività preparatorie per una valutazione.

Tuttavia, se il sistema di IA effettua profilazione di individui sarà sempre considerato ad alto rischio.

Quest'eccezione sarà fondamentale in quanto molti fornitori di sistemi di IA cercheranno di dimostrare che il loro sistema non costituisce un rischio, per evitare oneri normativi e costi elevati legati alla qualifica di IA ad alto rischio. Tuttavia, per godere di questa eccezione, il fornitore dovrà documentare la propria valutazione rispetto ad essa.

Anche se l'eccezione viene accolta, il sistema di IA dovrà comunque essere registrato nel database UE per i sistemi di IA ad alto rischio prima di essere commercializzato o utilizzato.

I fornitori di sistemi di IA ad alto rischio devono rispettare rigorosi requisiti per garantire l'affidabilità, la trasparenza e la responsabilità dei propri sistemi di IA. Tra gli obblighi, devono condurre valutazioni del rischio, utilizzare dati di alta qualità, documentare le decisioni tecniche ed etiche, monitorare le prestazioni del sistema, informare gli utenti sulla natura e lo scopo del sistema, consentire la supervisione e l'intervento umano e garantire l'accuratezza, la robustezza e la sicurezza informatica. Devono inoltre verificare la conformità dei loro sistemi alle normative prima di commercializzarli o metterli in servizio, e registrare i sistemi in un database dell'UE accessibile al pubblico.

La legge sull'IA impone obblighi rigorosi non solo al "fornitore" di un sistema di IA ad alto rischio, ma anche all'"importatore", al "distributore" e all'"utilizzatore" di tali sistemi.

Gli obblighi dell'importatore e del distributore riguardano principalmente la verifica della conformità dei sistemi di IA ad alto rischio che importano o distribuiscono. In generale, l'importatore deve verificare la conformità del sistema attraverso l'analisi di documentazione varia, mentre il distributore è tenuto a garantire la conformità CE (Conformité Européenne).

Pubblicato su: https://www.agenda17.it/

Anche l'"operatore", precedentemente noto come utente del sistema di IA, è soggetto a diversi obblighi quando impiega un sistema di IA ad alto rischio.

Un obbligo critico con potenziali impatti sul tema della responsabilità con il fornitore è che l'operatore deve utilizzare il sistema di IA ad alto rischio conformemente alle istruzioni d'uso del fornitore. Nel caso in cui un'azienda o i suoi clienti subiscano danni a causa dell'utilizzo di un sistema di IA ad alto rischio, il fornitore del sistema di IA probabilmente sosterrà che l'utente non ha seguito le istruzioni d'uso.

Inoltre, l'operatore ha l'obbligo di garantire, per quanto possibile, la presenza di supervisione umana e monitorare i dati di input e il funzionamento del sistema. Deve conservare i log automatizzati per almeno sei mesi.

Gli enti del settore pubblico e privato che forniscono servizi pubblici, come istruzione, sanità, alloggi, servizi sociali, nonché gli enti coinvolti nel credit scoring e nell'assicurazione sulla vita e sanitaria, sono tenuti a condurre una valutazione d'impatto sui diritti fondamentali prima di utilizzare un sistema di intelligenza artificiale ad alto rischio. Tale valutazione richiede che tali entità identifichino i rischi, le misure di supervisione, le misure di mitigazione del rischio, le categorie di persone fisiche coinvolte, la frequenza prevista d'uso e i processi operativi per i quali il sistema sarà impiegato.

3.6.Responsabilità e diritti lungo la catena del valore per i sistemi di IA ad alto rischio

Esiste un meccanismo simile alle norme sulla responsabilità per i danni da prodotti difettosi, in base al quale una parte diversa dal fornitore può essere considerata altresì fornitore. Importatori, distributori, rivenditori o qualsiasi terza parte possono essere considerati fornitori di un sistema di intelligenza artificiale ad alto rischio e saranno quindi soggetti a una serie di obblighi previsti dalla legge sull'AI, se una delle seguenti condizioni è soddisfatta:

- hanno apposto il proprio nome o marchio sull'apparecchiatura dopo che questa è stata già immessa sul mercato o messa in servizio;
- hanno apportato modifiche sostanziali dopo l'immissione sul mercato o messa in servizio, a condizione che il sistema resti ad alto rischio;
- hanno modificato lo scopo previsto del sistema di intelligenza artificiale, rendendolo ad alto rischio.

Inoltre, per diverso tempo si è dibattuto sulla questione se il GDPR conceda un diritto a una spiegazione all'interessato quando un titolare del trattamento adotta decisioni automatizzate individuali, inclusa la profilazione, che hanno effetti legali o simili per l'interessato.

La legge sull'IA ora conferma esplicitamente questo diritto, ma solo per i sistemi di intelligenza artificiale ad alto rischio elencati nell'allegato III: un individuo ha ora il diritto di ricevere spiegazioni significative sul ruolo del sistema di IA nel processo decisionale e sui principali elementi della decisione adottata.

In pratica si scatenerà una battaglia tra chi richiede spiegazioni e i fornitori che bloccano o limitano tali richieste sulla base di un segreto commerciale. Un buon esempio sono gli algoritmi di credit scoring, che costituiscono un sistema di intelligenza artificiale ad alto rischio.

Il modello di business e l'unico punto di forza di un'agenzia di credit scoring risiederanno nei pesi e nei parametri esatti utilizzati nel modello, che possono in gran parte essere protetti da segreti commerciali. È probabile che, nella pratica, si debba fare un esercizio di bilanciamento, seguendo il parere dell'avvocato generale Pikamäe in una recente causa (causa C-634/21) dinanzi alla Corte di giustizia dell'Unione europea: mentre, in linea di principio, la tutela dei segreti commerciali o della proprietà intellettuale costituisce un motivo legittimo per un'agenzia di informazioni creditizie di rifiutarsi di divulgare l'algoritmo utilizzato per calcolare il punteggio dell'interessato, non può in nessun caso giustificare un rifiuto assoluto di fornire informazioni, tanto più dove esistono mezzi di comunicazione adeguati che aiutino la comprensione garantendo al contempo un certo grado di riservatezza.

La legge sull'AI riconosce il diritto di presentare reclami presso un'autorità di vigilanza del mercato a qualsiasi persona fisica o giuridica che abbia ragioni di ritenere che la legge sull'AI sia stata violata. Si tratta di un ambito piuttosto esteso per esercitare tale diritto, poiché praticamente non sono richiesti requisiti particolari per farlo. Questo rappresenta una differenza significativa rispetto ad altri strumenti normativi, come il GDPR, in cui gli interessati possono presentare reclami solo se il trattamento dei loro dati personali è coinvolto.

Pubblicato su: https://www.agenda17.it/

3.7.I modelli di IA per scopi generali non sono sistemi di IA

I modelli di IA per scopi generali (GPAI) sono regolamentati e classificati specificamente dalla legge sull'AI. La legge sull'IA distingue tra gli obblighi applicabili a tutti i GPAI e gli obblighi aggiuntivi per i modelli GPAI con rischi sistemici.
Poiché i modelli sono regolamentati separatamente dai sistemi di intelligenza artificiale, un modello non sarà mai considerato un sistema di intelligenza artificiale ad alto rischio, in quanto non è un sistema di intelligenza artificiale.

Tuttavia, un sistema GPAI costruito su un modello GPAI potrebbe costituire un sistema di intelligenza artificiale ad alto rischio.

Il regolamento fissa una soglia per identificare i sistemi ad alto impatto, che hanno maggiori effetti sulla popolazione e perciò devono rispettare obblighi più stringenti.

Il valore, come dichiarato a dicembre, è un potere di calcolo pari a 10^25 FLOPs(*floating point operations per second)*, un'unità di misura della capacità computazionale). Al momento, solo GPT-4 diOpenAI, Gemini di Google e qualche modello cinese rispetterebbero questa caratteristica. Ma dovranno essere gli sviluppatori a comunicarlo alla Commissione, che per adesso non si esprime sui modelli già nei radar dell'AI Act e potrà intervenire se viene a sapere che un sistema ad alto impatto non si è dichiarato tale.
Da Bruxelles fanno sapere che la soglia potrà essere modificata in futuro, per meglio rispondere alle evoluzioni di mercato.

Le AI ad alto impatto dovranno applicare ex ante delle regole su sicurezza informatica, trasparenza dei processi di addestramento e condivisione della documentazione tecnica prima di arrivare sul mercato. Al di sotto si collocano tutti gli altri *foundational models*. Tra cui le due startup *made in Europe*: la francese Mistral e la tedesca Aleph Alpha. In questo caso l'AI Act scatta quando gli sviluppatori commercializzano i propri prodotti. E sono esclusi i modelli offerti con licenza open source.

In tal senso, i fornitori di modelli GPAI sono soggetti a obblighi distinti che possono essere considerati una versione semplificata degli obblighi per i sistemi di IA. Tra le altre cose, devono creare e mantenere la documentazione tecnica, elaborare una politica relativa al rispetto della normativa sul copyright e creare un riepilogo dettagliato del contenuto utilizzato per addestrare il modello GPAI.

Pubblicato su: https://www.agenda17.it/

I fornitori di modelli GPAI con rischi sistemici hanno ulteriori obblighi, tra cui lo svolgimento di valutazioni sui modelli, la valutazione e l'attenuazione dei rischi sistemici, la documentazione e la segnalazione di gravi incidenti all'Ufficio AI e alle autorità nazionali competenti, nonché garantire un'adeguata protezione della sicurezza informatica.

Essendo una terza categoria di sistemi di IA regolamentati (oltre alle pratiche di IA vietate e all'IA ad alto rischio), la legge sull'IA impone obblighi di trasparenza per quattro categorie di sistemi di IA e modelli GPAI:
- sistemi di IA destinati a interagire direttamente con persone fisiche (ad esempio assistenti AI);
- sistemi di intelligenza artificiale, inclusi i sistemi GPAI, che generano contenuti sintetici audio, immagini, video o testi (ad esempio Midjourney, DALL-E);
- sistemi di riconoscimento emotivo o categorizzazione biometrica (ad esempio ShareArt);
- deepfakes.

In questi casi l'utente dovrà essere informato sul sistema AI. In alcuni casi, il contenuto dovrà essere etichettato in modo leggibile dalla macchina in modo che possa essere identificato come contenuto generato o manipolato artificialmente.

La legge sull'intelligenza artificiale prevede eccezioni a tale obbligo in alcune circostanze per le forze dell'ordine o quando il sistema di intelligenza artificiale viene utilizzato per scopi artistici, satirici, creativi o simili.

4. Preoccupazioni relative ai diritti umani

È evidente che l'AI Act dia priorità all'economia, agli affari e all'innovazione rispetto alle preoccupazioni morali e che i diritti umani siano considerati solo marginalmente.

È interessante notare che i diritti fondamentali compaiono raramente nel testo principale; l'obiettivo principale dell'AI Act sembra essere maggiormente focalizzato sui mercati e sull'accesso al mercato dei sistemi di intelligenza artificiale.

Secondo Almada e Petit, l'AI Act deve essere considerata, innanzitutto, uno strumento di sicurezza del prodotto. Tale quadro, tuttavia, sembra essere inadatto a gestire le questioni relative ai diritti umani: i tentativi di affrontare le preoccupazioni relative ai diritti fondamentali attraverso una lente di sicurezza del prodotto potrebbero trascurare questioni cruciali come il potenziale indebolimento della fiducia sociale e della legittimità pubblica derivante da tecnologie come i chatbot e i deepfake.

Pubblicato su: https://www.agenda17.it/

Concentrandosi esclusivamente sui rischi e sulla sicurezza dei prodotti, il quadro dell'UE sembra trascurare altri rischi significativi, soprattutto per quanto riguarda i diritti fondamentali, che potrebbero essere messi in gioco dall'adozione e dall'utilizzo dei sistemi di intelligenza artificiale. In definitiva, l'AI Act potrebbe mettere in secondo piano la protezione dei diritti che non rientrano facilmente nella visione della sicurezza del prodotto, come i diritti politici.

Argomentazioni simili sono state avanzate da alcuni studiosi che hanno accusato la Commissione di concentrarsi sulla standardizzazione e armonizzazione del mercato unico invece che sui diritti umani.
Sembra che l'UE abbia scelto un approccio alla regolamentazione dell'IA che pone maggiore enfasi sulla definizione di standard e sull'influenza politica, sfruttando l'"effetto Bruxelles", mentre depriorizza la protezione e la promozione dei diritti umani.

Floridi aggiunge che, sebbene il fondamento etico dell'AI Act sia la tutela della dignità umana e dei diritti fondamentali, l'attuale proposta sembra essere più top-down, meno flessibile e meno focalizzata sulla protezione dei cittadini e dei loro diritti rispetto al GDPR e ad altre iniziative normative.

In sintesi, l'AI Act non fornisce una sufficiente protezione dei diritti fondamentali, non regolamenta adeguatamente la manipolazione permessa dall'intelligenza artificiale, il social scoring e l'uso di sistemi di riconoscimento biometrico. Inoltre, vieta l'uso di sistemi di riconoscimento delle emozioni e di categorizzazione biometrica remota in tempo reale o li classifica come ad alto rischio. Infine, il quadro normativo lascia troppa discrezionalità ai fornitori di IA a causa di una scarsa comprensione dei diritti fondamentali da parte dell'AI Act, incapace di garantire il necessario pilastro di un'IA giuridicamente affidabile.

I ricercatori e gli operatori sottolineano che la definizione di sistemi di intelligenza artificiale è eccessivamente ampia e manca di chiarezza, il che potrebbe portare a incertezza giuridica per sviluppatori, operatori e utenti di sistemi di intelligenza artificiale o, peggio ancora, a una regolamentazione inefficace.

È pertanto necessario rivedere o affinare la definizione di IA e chiarire il campo di applicazione dell'AI Act. Ad esempio, invece di parlare di sistemi di intelligenza artificiale, si potrebbe concentrarsi su sistemi decisionali automatizzati o algoritmici (ADS); contrariamente ai sistemi di intelligenza artificiale, l'ADS è considerato da molti la soluzione migliore, poiché è un termine indipendente dalla tecnologia e, quindi, a prova di futuro.

Pubblicato su: https://www.agenda17.it/

Alcuni autori sollecitano la Commissione a modificare di conseguenza la terminologia dell'AI Act. Inoltre, i critici chiedono che venga affrontata la "potenziale lacuna nella protezione giuridica relativa all'IA militare" e che venga chiarita "l'applicabilità dell'AI Act alle agenzie di sicurezza e di intelligence nazionali", poiché non è chiaro perché i sistemi di IA utilizzati dai militari e dalle agenzie di intelligence siano prevalentemente esentati dalle disposizioni dell'AI Act.

5. Critiche all'approccio basato sul rischio

Il quadro basato sul rischio dell'UE è stato criticato per due motivi: in primo luogo, i divieti dei sistemi di intelligenza artificiale (IA) non sono sufficientemente ambiziosi, e in secondo luogo, alcuni sistemi di IA elencati come a rischio basso o minimo hanno conseguenze sociali di vasta portata che potrebbero portare a gravi impatti sui diritti fondamentali e dovrebbero quindi essere considerati ad alto rischio.

Ad esempio, alcuni studiosi sostengono che i "sistemi manipolativi di intelligenza artificiale" esistenti, come i deepfake, e tutte le forme di "social scoring" dovrebbero essere vietate.

Essi sostengono inoltre che l'AI Act dovrebbe includere un divieto generale di sistemi di identificazione e categorizzazione biometrica, compresa la biometria per uso privato, nonché il divieto di sistemi di riconoscimento delle emozioni, della polizia predittiva, dell'uso di sistemi di intelligenza artificiale nel controllo delle frontiere e della migrazione e il divieto globale delle tecnologie di riconoscimento facciale (FRT).

Sulla base di queste osservazioni, durante la stesura del testo l'uso delle FRT in tempo reale negli spazi accessibili al pubblico per scopi di controllo del rispetto della legge è stato vietato, a meno che gli Stati membri dell'Unione Europea decidano di autorizzarle per importanti motivi di sicurezza pubblica e siano fornite adeguate autorizzazioni amministrative o giudiziarie.

Ciò implica che una vasta gamma di FRT utilizzati per scopi diversi dal controllo del rispetto della legge (ad esempio, gestione delle frontiere, trasporto pubblico, mercati, ecc.) potrebbe essere autorizzata se superano la valutazione di conformità e rispettano i relativi requisiti di sicurezza prima di entrare nel mercato dell'UE.

Pertanto, l'impiego di sistemi di riconoscimento facciale e biometrico in tempo reale è una applicazione proibita, poiché, come si legge nelle premesse, può portare "a risultati marcati da pregiudizi e provocare effetti discriminatori", ad eccezione di tre "situazioni ampiamente elencate e ben definite", nelle quali il ricorso al riconoscimento facciale "è necessario per raggiungere un sostanziale pubblico interesse, la cui importanza supera i rischi".

Pubblicato su: https://www.agenda17.it/

E i tre casi derogativi sono la ricerca di vittime di reati e di persone scomparse, le minacce certe alla vita o alla sicurezza fisica delle persone o di attacco terroristico e la localizzazione e l'identificazione dei presunti autori di una lista di 16 reati contenuti nell'allegato IIa.

L'elenco comprende: terrorismo; traffico di esseri umani; abusi sessuali su minori e pedopornografia; traffico di droghe e sostanze psicotrope; traffico illecito di armi, munizioni ed esplosivi; omicidio o gravi feriti; traffico di organi; traffico di materiale radioattivo e nucleare; sequestro di persona e ostaggi; crimini sotto la giurisdizione della Corte penale internazionale; dirottamento di aerei e navi; stupri; crimini ambientali; rapine organizzate e armate; sabotaggio; partecipazione a una organizzazione criminale coinvolta in uno o più crimini tra quelli elencati.

Il riconoscimento biometrico da remoto in tempo reale dovrebbe essere utilizzato "solo per confermare l'identità" della persona individuata come target, dopo aver valutato il rischio di non utilizzare questa tecnologia rispetto ai risultati che si possono ottenere, e solo quando strettamente necessario "nello spazio e nel tempo".

Le forze di polizia devono valutare l'impatto sui diritti fondamentali dei cittadini prima di adottare questi strumenti e ottenere l'autorizzazione da un giudice o da un ente indipendente.

L'AI Act prevede una procedura di emergenza che consente l'attivazione della sorveglianza biometrica, con 24 ore di tempo per richiedere l'autorizzazione. In caso di mancata autorizzazione, l'uso del riconoscimento facciale deve essere bloccato immediatamente e tutti i dati cancellati.

I garanti nazionali dei dati personali e del mercato devono inviare annualmente alla Commissione un rapporto sull'uso dei sistemi di riconoscimento biometrico in tempo reale, compresi eventuali usi non autorizzati.

Gli Stati dell'Unione possono adottare leggi nazionali per ampliare il campo di applicazione della sorveglianza biometrica, nel rispetto dei limiti stabiliti dall'AI Act.

Le stesse regole si applicano anche al riconoscimento facciale utilizzato ex post, con una finestra di 48 ore per ottenere l'autorizzazione in caso di emergenza.

Tuttavia, l'approccio basato sul rischio dell'UE ha alcuni difetti, secondo i critici. In primo luogo, molti sistemi ad alto rischio potrebbero essere consentiti e legittimati dall'AI Act, anche se non sono stati adeguatamente testati e dibattuti pubblicamente.

Pubblicato su: https://www.agenda17.it/

In secondo luogo, non è chiaro quando si raggiunge la soglia dell'alto rischio, ovvero quando i sistemi sono considerati ad alto rischio rispetto a un rischio basso o minimo.

In terzo luogo, l'elenco dei sistemi ad alto rischio deve essere modificabile per rendere l'AI Act preparata al futuro.
Si osserva che l'articolo 5 dell'AI Act non può essere modificato, a differenza dell'articolo 7, il che significa che non è possibile aggiungere pratiche vietate una volta che l'AI Act è stata promulgata. Questo dovrebbe essere modificato e mentre si rivede l'elenco dei sistemi di intelligenza artificiale ad alto rischio, al pubblico dovrebbero essere fornite opportunità di consultazione e partecipazione.

Inoltre, i sistemi di intelligenza artificiale non classificati come ad alto rischio sono soggetti a una sotto-regolamentazione. Questi sistemi potrebbero avere impatti gravi e dannosi su individui e società, ma sono soggetti solo a un limitato insieme di requisiti di trasparenza.
Ad esempio, i sistemi che interagiscono con gli esseri umani (come i chatbot), i sistemi di riconoscimento delle emozioni, i sistemi di categorizzazione biometrica, compresi i sistemi di riconoscimento facciale, e i sistemi di intelligenza artificiale che generano o manipolano contenuti di immagini, audio o video (come i deepfakes) sono soggetti solo a un limitato insieme di requisiti di trasparenza.

In breve, l'AI Act lascia una vasta gamma di sistemi di intelligenza artificiale non regolamentati, nonostante possano avere un impatto grave sui diritti fondamentali.

6. Le sfide poste dall'IA generativa

L'AI Act potrebbe non essere in grado di affrontare adeguatamente le sfide poste dall'IA generativa per scopi generali. Si noti che il progetto di proposta della Commissione non includeva i termini "IA generativa" e "AI per scopi generali", e le parole "chatbot" e "deepfake" venivano menzionate solo una volta.

L'AI Act sembra presumere che i legislatori possano specificare in modo previdente una particolare funzionalità o caso d'uso dei sistemi di intelligenza artificiale. Tuttavia, come dimostra l'evoluzione recente dell'intelligenza artificiale generativa (ad esempio, grandi modelli linguistici e chatbot), è difficile prevedere in anticipo le funzionalità e i casi d'uso che potrebbero sollevare problemi di sicurezza e altre questioni etiche.

Pubblicato su: https://www.agenda17.it/

I problemi legati all'evoluzione tecnologica e alle relative incertezze pongono sfide significative per la regolamentazione dell'IA in generale e per l'AI Act in particolare.

Analoghe difficoltà sorgono anche nel contesto della standardizzazione dei sistemi di intelligenza artificiale: a causa del rapido cambiamento nella ricerca e nello sviluppo, gli standard potrebbero diventare obsoleti o richiedere regolamenti frequenti.

Uno dei principali problemi dell'AI Act riguarda il processo di autovalutazione e standardizzazione. Se il progetto di proposta venisse attuato così com'è, il processo di governance sarebbe principalmente basato sull'autocontrollo da parte dei fornitori, in particolare attraverso la cosiddetta "procedura di valutazione della conformità ex ante".

La sorveglianza di mercato da parte delle autorità degli Stati membri avviene solo a posteriori, non preventivamente. Questo processo viene eseguito dal fornitore stesso, e non da un terzo indipendente, esterno.

Di conseguenza, l'approccio proposto lascia la valutazione preliminare del rischio, compresa l'identificazione dei sistemi di intelligenza artificiale ad alto rischio, al fornitore e allo sviluppatore, concedendo loro un ampio margine di discrezionalità.

Ad esempio, possono decidere se il software utilizzato è un sistema di intelligenza artificiale, se il sistema può causare danni e come soddisfare i requisiti obbligatori del Titolo III, Capitolo 2 dell'AI Act; inoltre, teoricamente possono classificare le tecnologie ad alto rischio come conformi alle regole tramite la procedura di autovalutazione.

I critici temono che le norme sulle pratiche di IA vietate e ad alto rischio possano rivelarsi inefficaci poiché la valutazione del rischio è lasciata al fornitore/sviluppatore.

Inoltre, oltre a concedere un eccessivo potere discrezionale ai fornitori di intelligenza artificiale, l'AI Act permette anche un margine significativo di intervento alla "standardizzazione" e agli "organismi notificati" per quanto riguarda l'applicazione e l'attuazione dell'AI Act e la decisione su altre questioni politiche ed etiche delicate.

Il processo di standardizzazione è fondamentale per l'autovalutazione della conformità e richiede ai fornitori di seguire "standard armonizzati" sviluppati da organizzazioni europee di standardizzazione come il CEN (Comitato europeo di standardizzazione) o il CENELEC (Comitato europeo di standardizzazione elettrotecnica).

Pubblicato su: https://www.agenda17.it/

Tuttavia, delegare il potere di regolamentazione e standardizzazione a tali organismi, cioè trasferire la competenza pubblica di regolamentazione ad associazioni private, solleva diverse questioni etiche, come il rischio di esautorazione normativa.

Gli intermediari come gli organismi notificati potrebbero non avere l'indipendenza necessaria e potrebbero essere influenzati dagli interessi dei fornitori di IA, creando potenziali conflitti di interesse; inoltre, i grandi fornitori e sviluppatori di IA (le "big tech") detengono un significativo potere di mercato, aumentando i rischi di deregolazione.

Complessivamente, i ricercatori temono che l'odierna attenzione all'autovalutazione e all'autoregolamentazione aziendale sia insufficiente e possa portare a una regolamentazione inefficace o insufficiente. Questo rischio di elusione della regolamentazione è ulteriormente aggravato dalla dipendenza dall'autovalutazione dell'AI Act, che oltre alle autovalutazioni, utilizza anche il soft law e i codici di condotta per i sistemi di intelligenza artificiale a basso o minimo rischio.

7. Asimmetrie di potere e discriminazioni algoritmiche

La delega o l'esternalizzazione del potere normativo agli organismi europei di standardizzazione (ossia organismi di standardizzazione e notificati) potrebbe portare a una mancanza di controllo democratico e impedire alle parti interessate, come le organizzazioni della società civile e i gruppi di difesa dei consumatori, di influenzare il processo decisionale, nonché di disporre di mezzi giuridici per monitorare e controllare le norme una volta adottate.

La mancanza di partecipazione delle parti interessate si manifesta nei processi di standardizzazione, dove spesso manca un coinvolgimento adeguato di tutti i gruppi di parti interessate rilevanti, inclusi i rappresentanti dei consumatori e della società civile.

In particolare, le organizzazioni dei consumatori spesso incontrano difficoltà a partecipare a tali processi a causa di limiti finanziari, competenze, esperienza o rappresentanza politica, risultando in esclusione di alcuni gruppi di portatori d'interesse e mancanza di legittimità democratica.

Le asimmetrie di potere emergono anche nello sviluppo degli standard, dove i rappresentanti dell'industria hanno una significativa influenza rispetto alle voci della società civile e dei difensori dei consumatori, che spesso non hanno una rappresentanza a livello dell'UE.

Inoltre, nel contesto dell'intelligenza artificiale si osservano squilibri di potere tra fornitori e sviluppatori di IA da un lato e altri gruppi di stakeholder e organismi di standardizzazione e notificati dall'altro.

La mancanza di trasparenza sulle attività degli organismi di standardizzazione e notificati è spesso dovuta alla terziarizzazione delle attività di test, ispezione e certificazione a terzi, con critiche per l'eccessiva fiducia nelle autovalutazioni aziendali anziché nel monitoraggio indipendente e imparziale.

I problemi di ricerca del consenso riguardano la complessità normativa nell'affrontare questioni come l'equità algoritmica, l'uguaglianza e la trasparenza nell'addestramento dei dati. La capacità del CEN o del CENELEC di affrontare tali questioni è dibattuta, con la previsione di difficoltà nel negoziare e trovare un consenso tra i diversi gruppi di interesse durante il processo di standardizzazione.

Gli autori sottolineano la necessità che tali questioni siano affrontate dalle istituzioni democraticamente responsabili come la Commissione e il Parlamento, anziché essere delegate agli organismi di standardizzazione non eletti democraticamente.

Le persone negativamente colpite dai sistemi di intelligenza artificiale hanno il diritto di presentare reclamo alle autorità di vigilanza del mercato, come previsto dal GDPR. Tuttavia, il diritto a presentare reclami alle autorità di vigilanza del mercato potrebbe mancare di efficacia, poiché non è chiaro con quale tempestività e determinazione tali autorità garantiscono il rispetto dell'adesione e la responsabilità degli infrattori.

Inoltre, il diritto a una spiegazione sui processi decisionali, soprattutto per i sistemi di intelligenza artificiale classificati ad alto rischio, solleva dubbi riguardo alla praticità e all'accessibilità nell'ottenere spiegazioni significative dagli operatori. L' efficacia di tali meccanismi rimane incerta, considerando l'assenza di disposizioni come il diritto alla rappresentanza delle persone fisiche o la possibilità per le organizzazioni di interesse pubblico di presentare reclami alle autorità di vigilanza nazionali.

L'uso di sistemi di intelligenza artificiale nei processi di gestione delle risorse umane solleva questioni etiche legate alla privacy, alla protezione dei dati, alla discriminazione e al monitoraggio costante. Questo monitoraggio potrebbe causare danni psicologici, stress e pressioni sul lavoro, influenzando negativamente la produttività e limitando i diritti dei lavoratori.

Pubblicato su: https://www.agenda17.it/

I critici sottolineano che l'AI Act sembra concentrarsi meno sulla protezione dei diritti dei cittadini, inclusi quelli dei lavoratori, rispetto al GDPR e ad altre leggi dell'UE. L'elenco delle pratiche vietate di intelligenza artificiale sul lavoro è limitato e i processi di supervisione e valutazione non coinvolgono ancora i lavoratori o le parti sociali, lasciando i lavoratori privi di protezione adeguata contro i rischi dovuti al monitoraggio basato sull'intelligenza artificiale sul luogo di lavoro e alle intrusioni nella privacy.

Il finanziamento adeguato e il personale sufficiente sono cruciali per far rispettare l'AI Act, monitorare la conformità e sanzionare le violazioni. La Commissione stima che l'attuazione dell'AIA richieda da uno a venticinque dipendenti aggiuntivi a tempo pieno per ogni Stato membro, ma alcuni studiosi ritengono che ciò non sia sufficiente per garantire un efficace regime di governance dell'IA.

Un problema correlato riguarda la mancanza di conoscenze tecniche e competenze legali nelle organizzazioni di definizione degli standard e nelle autorità di vigilanza del mercato rispetto ai fornitori/sviluppatori di IA. Questa mancanza di competenze porta spesso a una dipendenza da agenti di mercato e all'outsourcing di compiti e responsabilità a attori privati, aggravando i problemi di trasparenza, responsabilità e legittimità democratica.

Pertanto, i critici esortano a fornire risorse finanziarie migliori e a incrementare il personale delle agenzie di regolamentazione, assumendo ad esempio analisti di dati, specialisti di intelligenza artificiale e ricercatori indipendenti.

Per quanto riguarda le ambiguità istituzionali, il coinvolgimento di sette istituzioni di vigilanza nell'attuazione dell'AI Act solleva critiche sulla mancanza di coordinamento e chiarezza sui rapporti tra di esse.
Ancora non è chiaro come la supervisione dei regimi di governance già esistenti, come il GDPR, il DSA e il DMA, si integrerà con il nuovo regolamento sull'IA e con il Comitato europeo per l'intelligenza artificiale (EAIB).

L'AI Act non affronta in modo adeguato i rischi di sostenibilità derivanti dai sistemi di intelligenza artificiale, come il consumo energetico elevato e le emissioni di gas serra dei data center. Non ci sono indicazioni dirette sulle questioni ambientali, come il cambiamento climatico, la sostenibilità ecologica e l'intelligenza artificiale verde. Questo solleva preoccupazioni sulla priorità data alle esigenze umane rispetto alla tutela dell'ambiente secondo un approccio antropocentrico seguito dall'UE.

Pubblicato su: https://www.agenda17.it/

Altri aspetti critici riguardano la valutazione del rischio e dell'impatto, la supervisione umana, la parzialità e l'equità algoritmica e i diritti di proprietà intellettuale. Alcuni esperti sostengono che sia necessaria una valutazione del rischio dell'IA, in particolare riguardo al potenziale discriminatorio dei sistemi di intelligenza artificiale, e una valutazione dell'impatto sui diritti umani. Pertanto, esortano i legislatori a garantire i "diritti umani per impostazione predefinita" o una "progettazione dei diritti umani", simile alla privacy per impostazione predefinita.

I critici evidenziano che l'AI Act non chiarisce quando, come e in quale fase è richiesta la supervisione umana, il che porta a una notevole mancanza di chiarezza e certezza.

Inoltre, l'AI Act non specifica la nozione di bias algoritmico e di equità, né indica quali forme di pregiudizio sono vietate e come dovrebbero essere mitigate. Non richiede neanche a sviluppatori, operatori e utenti di fornire informazioni a coloro che sono influenzati negativamente dalle previsioni o decisioni basate sull'IA.

La mancanza di menzione sulla non discriminazione e sull'equità solleva ulteriori preoccupazioni. Gli esperti suggeriscono il coinvolgimento dei centri nazionali di competenza sull'IA e delle organizzazioni per i diritti umani per fornire input qualitativi e proposte per valutare i bias nei sistemi di IA.
Infine, è importante stabilire linee guida per garantire la trasparenza e l'informazione, preservando nel contempo i diritti di proprietà intellettuale e i segreti commerciali. La questione su quanto il codice sorgente debba essere divulgato alle autorità rimane essenziale.

8. Conclusione

In conclusione, l'analisi dell'AI Act ha evidenziato punti di forza e debolezza, con proposte di riforma per affrontare le sfide e rafforzare il regolamento.

Il documento propone diverse misure, come il divieto di determinati sistemi di IA ad alto rischio utilizzati sia da attori pubblici che privati e la classificazione di altri sistemi come ad alto rischio.

Per colmare le lacune, è necessario eliminare le eccezioni elencate nell'art. 5, garantire la modificabilità dell'elenco delle pratiche di IA vietate e dei sistemi ad alto rischio, introdurre valutazioni obbligatorie da parte di terzi anziché autovalutazioni, rafforzare il controllo democratico (e giudiziario) del processo di standardizzazione, inclusi diritti di consultazione e l'inclusione di tutte le parti interessate attraverso un dialogo globale.

Pubblicato su: https://www.agenda17.it/

È anche necessario proteggere meglio i diritti fondamentali con valutazioni obbligatorie del rischio e dell'impatto sui diritti umani per tutti i sistemi basati sull'intelligenza artificiale, incorporare requisiti relativi alla sostenibilità per i fornitori di IA, compresi riferimenti espliciti al consumo significativo di risorse e all'IA verde e sostenibile, modificare l'onere della prova con il concetto di "illegittimità per impostazione predefinita" per obbligare gli sviluppatori a dimostrare la non creazione di danni prima del permesso sul mercato UE.

Inoltre, è necessario introdurre diritti di informazione e meccanismi di reclamo, rimedio e risarcimento individuali e collettivi, migliorare il coinvolgimento e la protezione dei lavoratori e dei sindacati, attraverso partecipazione, codeterminazione, diritti di segnalazione e meccanismi di denuncia, chiarire i ruoli e le responsabilità degli organismi di monitoraggio ed esecuzione e il loro rapporto tra loro e con altre istituzioni e normative dell'UE, rafforzare l'autonomia politica del Comitato Europeo per l'intelligenza artificiale (EAIB) e dotarla di poteri investigativi e normativi sufficienti, garantire che la banca dati dell'UE per i sistemi di IA includa tutti i sistemi, non solo quelli ad alto rischio, e che i dati siano accessibili al pubblico in un formato comprensibile, fornire maggiori finanziamenti e personale alle autorità di vigilanza del mercato a livello UE e degli Stati membri e lavorare per armonizzare gli standard e le linee guida internazionali sull'IA.

Pubblicato su: https://www.agenda17.it/