Intelligenza Artificiale e novità normative



Negli ultimi anni l'Intelligenza Artificiale si è imposta al centro del dibattito globale: la sua rapidissima evoluzione, col conseguente aumento dei potenziali campi di utilizzo, ha reso questa tecnologia – in particolare nella più recente accezione "generativa" – protagonista di accese discussioni.

La lettera in cui i CEO delle principali aziende sul mercato proponevano una "moratoria" per valutare le ricadute etiche dell'AI o le dimissioni di Geoffrey Hinton, considerato uno dei padri nobili dell'Intelligenza Artificiale, sono solo i passaggi più noti di una crescente tendenza che vede anche professionisti ed esperti assumere posizioni apertamente critiche rispetto ai rischi di un suo sviluppo incontrollato.

AI e cybersecurity: la posizione degli Enti di settore

L'ENISA Threat Landscape (ETL) 2023, undicesima edizione del rapporto annuale a cura dell'Agenzia europea per la Cybersecurity, include l'Intelligenza Artificiale fra i *key trends* più rilevanti entro lo scenario di minaccia attuale.

Il rapporto cita in particolare il fenomeno *deepfake* e l'aumento della disinformazione online, evidenziando l'impatto di *"generative artificial intelligence chatbots"* che stanno trasformando *"the way in which we work, live and play, all built around data sharing and analysis […] becoming a preferred target for cybercriminals as they are very susceptible to data poisoning".*

Analogamente, la recente *Roadmap for Artificial Intelligence 2023/24* pubblicata dalla statunitense Cybersecurity and

Infrastructure Security Agency (CISA) pone le sfide di sicurezza come necessario punto di partenza per un uso responsabile degli strumenti AI.

Secondo la direttrice dell'Agenzia, gli obiettivi del documento sono:

- promuovere l'impiego dell'AI per ampliare le risorse di cybersecurity;
- assicurare che i sistemi AI siano protetti dalle minacce informatiche;
- prevenire l'uso malevolo dell'AI contro le infrastrutture critiche.

La *roadmap* è parte dello sforzo per una *governance* federale dell'AI portato avanti dall'amministrazione Biden, che aveva già previsto diversi passaggi normativi (l'ultimo dei quali è l'*Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*).

Simili preoccupazioni hanno dato impulso alle numerose iniziative, portate avanti da Stati e organizzazioni internazionali, per definire *framework* tesi ad accompagnare la trasformazione operata dagli strumenti AI a partire dalla gestione dei relativi rischi.

Le nuove regole internazionali e l'approccio *risk-based*

Un esempio di tale approccio è rappresentato dai *"Principi Guida sull'Intelligenza Artificiale generativa"* elaborati nel corso del vertice G7 tenutosi a Hiroshima lo scorso ottobre, al fine di fornire un codice di condotta volontario per chiunque contribuisca allo sviluppo di sistemi AI avanzati.

Nel dettaglio, i progetti di principi guida chiedono alle organizzazioni di:

1. prevedere adeguate misure di identificazione,

valutazione e mitigazione del rischio lungo tutti i processi di sviluppo dei sistemi di AI più avanzati;
2. individuare le vulnerabilità, incidenti e possibili schemi di uso malevolo degli strumenti lanciati sul mercato;
3. rendere pubbliche le capacità, i limiti e gli ambiti applicativi (leciti e non) dei sistemi per garantire una sufficiente trasparenza;
4. garantire la condivisione d'informazioni e il *reporting* degli incidenti di sicurezza tra le organizzazioni che sviluppano tali sistemi, includendo industrie, governi, società civile e accademia;
5. sviluppare, implementare e rendere pubbliche le policy relative alla *governance* degli strumenti AI e alla gestione dei relativi pericoli, secondo un approccio *risk-based*;
6. investire in robusti controlli di sicurezza (includendo quella fisica e cyber, nonché ogni potenziale *insider threat*) e implementarli lungo l'intero ciclo di vita delle tecnologie;
7. prevedere solidi meccanismi (es. *watermarking*) per consentire agli utenti di identificare i contenuti *AI-generated*;
8. condurre ricerche tese a mitigare ogni *"societal, safety and security risk"*, investendo in concrete misure di mitigazione dei rischi così individuati;
9. dare priorità, nello sviluppo degli strumenti AI, alle principali sfide globali attuali, come il cambiamento climatico o l'accesso alla salute e all'istruzione;
10. procedere nello sviluppo di standard tecnologici condivisi a livello internazionale;
11. implementare adeguate misure a protezione dei dati personali e della proprietà intellettuale.

Il testo è attualmente oggetto di una *survey* della Commissione europea, i cui esiti contribuiranno a orientare la posizione dell'UE in seno al G7.

Nella medesima scia sembra già voler procedere il Regolamento AI in discussione presso il Parlamento europeo, ove si definisce *"ad alto rischio"* ogni strumento che – in ragione dello specifico campo di applicazione – possa avere un impatto su aspetti rilevanti della vita pubblica (è il caso del *Predictive Policing*) o privata (come il diritto alla salute).

Il futuro dell'AI è nella trasparenza

Resta da sperare che alle dichiarazioni formali seguano processi aperti e realmente inclusivi in cui siano coinvolti tutti gli *stakeholder* interessati allo sviluppo di strumenti AI sicuri, trasparenti e rispettosi dei diritti umani.

È quanto richiesto dall'appello con cui quasi 2000 personalità del comparto tecnologico, sotto l'egida di Mozilla, chiedono di considerare apertura, trasparenza ed eliminazione delle barriere all'accesso del mondo AI come priorità globali: *"When it comes to AI safety and security, openness is an antidote, not a poison"*.

A cura della Redazione

Condividi sui Social Network:
https://www.ictsecuritymagazine.com/notizie/intelligenza-artificiale-e-novita-normative/