Approvata la bozza dell'AI Act europeo: è un primo passo, ma servono regole internazionali per proteggerci da una possibile escalation

L'Unione europea accelera verso una normativa che, però, ha bisogno di essere condivisa per essere davvero efficace: al di là degli allarmismi inutili, è vero che l'AI presenta dei rischi importanti: da un lato, perché viene gestita da una manciata di società ed usata già oggi per i processi che riguardano la salute, il lavoro, la burocrazia, e persino la giustizia. Dall'altro, perché l'AI si sviluppa rapidamente e può arrivare a livelli di complessità che probabilmente non sapremo districare. Il rischio è non sapere perché l'AI prenda certe scelte e affidarci a lei ciecamente

Di Shalini Kurapati, Co-Founder e CEO di Clearbox AI

L'Unione europea sta cercando di regolamentare l'intelligenza artificiale. Un esercizio complesso perché se da un lato bisogna proteggere le persone dai rischi di una tecnologia così potente e complessa, dall'altro non si può – e non si deve – fermare la possibilità di migliorare diversi ambiti della nostra vita quotidiana. A complicare il tentativo dell'Ue, però, è anche la portata di una norma che, per essere efficace, deve essere riconosciuta a livello globale. Altrimenti rischia di essere agevolmente aggirata. Infatti le principali società che sviluppano l'AI nel mondo sono oggi Microsoft/Open AI, Google, Meta, Tesla, solo per citarne alcune: un panorama che coinvolge gli Stati Uniti, ma che inizia a coinvolgere anche la Cina. E gli Stati Uniti, per il momento, non hanno alcuna intenzione di frenare la ricerca e lo sviluppo: per quanto l'amministrazione Biden sia consapevole dei rischi, dopo tanti discorsi tutto ciò che è stato fatto ad oggi – di fatto – sono una serie di misure *non legislative* che suggeriscono ai colossi dell'AI come dovrebbero comportarsi per auto-regolare la propria tecnologia.

Normativa EU tra soluzioni e problematicità

L'idea del Vecchio continente è quella di classificare non tanto gli algoritmi, quanto l'intensità del rischio relativo al loro utilizzo: basso, medio, alto, inaccettabile. È un primo passo molto utile che serve a vietare quelle tecnologie che hanno un rischio troppo elevato in termini di violazione dei diritti fondamentali delle persone. Se l'AI viene usata per bloccare lo spam, il rischio è basso, se invece viene usata per la diagnosi e l'analisi delle forme tumorali è evidente che il rischio è molto più alto (può sbagliare, può discriminare, ecc). Senza dimenticare la possibilità che l'AI venga usata per il *social scoring*, ovvero per classificare la reputazione delle persone (come fa la Cina per i propri cittadini) - e il rischio per questo utilizzo sarebbe classificato come "inaccettabile".

È una norma che però presenta almeno una problematicità: dal momento che l'intelligenza artificiale si sviluppa rapidamente, non è sempre facile capire in anticipo se un determinato tipo di software sia pericoloso o meno. Un esempio su tutti: il caso dei chatbot. Se fino a un paio di anni fa sarebbero stati classificati come algoritmi a basso rischio, oggi con Chat-GPT lo scenario è radicalmente cambiato.

D’altra parte non esisterà mai una legge perfetta e qualunque norma approvata dovrà essere – in qualche modo – flessibile. Capace di adattarsi, o essere adattata, ai cambiamenti della società e agli sviluppi della tecnologia.

Regole globali: l’unica via davvero efficace

Forse ancora più importante di avere una regolamentazione flessibile è l’urgenza di avere delle norme condivise. Infatti, per quanto l’AI Act possa regolare i software del vecchio continente, non può farlo per quelli che vengono sviluppati in altre nazioni e questo, in un mondo globalizzato, è chiaramente un problema per tutti.

La proposta di raggiungere una cooperazione globale non è nuova: ci sono state delle prime discussioni nel G7 e piani per includere Paesi ricettivi come India, Indonesia e Brasile per andare verso un accordo globale sulla regolamentazione AI e protocolli di sicurezza, ma non abbastanza velocemente. Sono solo piccoli passi, almeno per ora.

Se gli allarmi più o meno esagerati sull’impatto disastroso che l’intelligenza artificiale avrebbe sul mondo si rincorrono da anni, è però vero che, senza scendere in scenari distopici, è ormai prioritario trovare un’intesa su larga scala globale, esattamente come dovrebbe accadere per proteggere il clima.

Verso un’AI più giusta

La questione è ampia e complessa, certo, ma le soluzioni possibili per fare in modo che ci sia un controllo etico dell’intelligenza artificiale sono molte: una delle proposte contenute nella bozza dell’AI Act è, per esempio, quella di creare dei processi che si servano dei **dati sintetici** (che è poi quello che facciamo in Clearbox AI). Questo aspetto è significativo perché solitamente le regolamentazioni non fanno riferimento diretto a tecniche specifiche. I dati sintetici sono costruiti partendo da dati reali e generandone di nuovi che però mantengono le proprietà statistiche di quelli originali: quindi da un lato permettono di anonimizzare i risultati (quindi non violare la privacy) e dall’altro anche di agire sulla riduzione dei bias. Anche per questo nell'AI Act c'è molto focus sul dato che sta alla base dell’intelligenza artificiale, la sua qualità e la sua validazione in termini di robustezza e rappresentazione, oltre ai requisiti di *data privacy* e *data minimization*. I dati sintetici hanno il potenziale per giocare un ruolo fondamentale nella realizzazione di questi requisiti verso un'AI sicura e affidabile.

Un’altra strada percorribile è quella di affidare a un ente terzo l’obbligo di certificare periodicamente lo stato dell’arte.

Insomma, L'AI Act europeo è l’apripista nell'istituzione di regole esaustive e applicabili per garantire la sicurezza dell'IA e l'innovazione responsabile. Nonostante alcuni limiti e sebbene potrebbe richiedere da 2 a 3 anni per entrare in vigore (ma la Commissione Europea sta lavorando a delle linee guida che le aziende possono iniziare ad adottare nel frattempo) ci sono comunque varie ragioni per sostenerlo fortemente, prima tra tutte perché ha creato un punto di riferimento che può essere utile per la cooperazione nella regolamentazione globale dell'intelligenza artificiale. E una cosa è certa: la finestra temporale per la normazione dell’intelligenza artificiale si sta riducendo rapidamente. Non ci sarà alcuna esplosione nucleare, ma è una come una palla di neve che rotolando verso valle diventa sempre più grande e sempre più difficile da fermare. E’ per questo che è importante agire adesso, come sta già facendo l’Europa, ma a livello mondiale.

Informazioni su Clearbox AI

Clearbox AI è la startup tech italiana che aiuta le aziende a lanciare progetti di AI e di Analytics attraverso la generazione di dati sintetici di alta qualità. La missione aziendale consiste nel comprendere e risolvere le sfide che le imprese incontrano nello sviluppo dei processi di Intelligenza Artificiale. Questi ostacoli sono spesso legati ai dati sensibili che sono difficili da gestire a livello di privacy e che possono non essere abbastanza rappresentativi per tutte le fasce di popolazione, o la loro quantità non è sufficiente per garantire risultati di successo. Il Data Engine di Clearbox AI è fondato su tecnologia proprietaria e agnostica basata su modelli generativi avanzati, creata anche grazie alle solide radici nel mondo della ricerca del team. La soluzione supporta qualsiasi tipo di azienda sia per incrementare la disponibilità di dati, la loro qualità e quindi mitigando eventuali bias interni ai dati, ma anche per aiutarle in termini di policy, compliance e privacy degli stessi.

Clearbox AI è stata recentemente selezionata dalla Commissione Europea per il progetto **Women TechEU**, che supporta le startup deep-tech guidate da donne per valorizzare il talento e favorire l'innovazione nell'ecosistema tech europeo. È inoltre vincitrice del Premio Nazionale dell'Innovazione (categoria ICT, 2019), il Seal of Excellence dell'Unione Europea, ed è stata selezionata da **Fortune Italia** come una delle migliori startup AI del paese.

Ufficio stampa Clearbox AI

ddl studio: innovationteam@ddlstudio.net
Mara Linda Degiovanni: +39 3496224812
Elisa Giuliana: +39 3386027361