Come supportare un approccio responsabile all' Intelligenza Artificiale

Francesca Rossi, IBM
Indagine Conoscitiva sull'IA, 3 Ottobre 2023

Vorrei ringraziare la Commissione per l'opportunita' di intervenire all'interno dell'indagine conoscitiva sull'IA e portare la mia esperienza a supporto dell'approfondimento avviato in ambito Parlamentare in Italia su un tema cosi' pervasivo e di impatto sulla societa'.

Vorrei brevemente presentarmi. Dopo un dottorato in Informatica all'Universita' di Pisa, ho contribuito alla ricerca in IA per circa 35 anni, 25 dei quali nelle universita' Italiane, a Pisa e poi a Padova, e poi in IBM negli Stati Uniti. In IBM, oltre a guidare progetti di ricerca, ho anche il ruolo di leader globale per l'etica dell'IA, con la responsabilita' di guidare le attivita' dell'azienda nell'ambito dell'etica dell'IA, dai principi alla loro implementazione in azioni concrete in tutte le divisioni dell'azienda, incluse attivita' di educazione, linee guida per gli sviluppatori di IA, processi di valutazione e gestione del rischio, e governance interna, attuata tramite il comitato per l'etica dell'IA, e anche attraverso le molte partnership con le universita', altre aziende, e organizzazioni globali. Sono un membro del board della Partnership on AI, dello Steering Committee della Global Partnership on AI (di cui fa parte anche l'Italia), sono co-chair del gruppo di esperti OECD sul futuro dell'IA, e ho fatto parte del Gruppo di esperti di alto livello sull'IA della Commissione Europea. Attualmente, sono anche il presidente di AAAI, l'associazione mondiali dei ricercatori in IA.

Benefici e rischi dell'IA.

Come tutti sappiamo, l'IA e' una scienza e una tecnologia con abilita' sorprendenti ma che introduce anche rischi e sfide significative. L'IA non e' una nuova tecnologia: viene usata gia' da molti anni in molti aspetti della nostra vita. La sua evoluzione e' iniziata con approcci cosiddetti "simbolici" e basati su regole, dove indicavamo alle macchine i passi da seguire per risolvere un problema. Questo approccio funzionava bene in molti scenari ma era troppo poco flessibile per funzionare bene al di fuori di contesti molto controllati. Siamo quindi passati ad approcci basati su dati e machine learning, in cui le macchine apprendono come risolvere un problema dall'analisi di dati su problemi simili. Questo ha reso l'IA piu' capace di interpretare l'ambiente in cui viene usata e quindi di prendere decisioni migliori e piu' personalizzate. Ma ha anche introdotto considerazioni legate alla privacy e la governance dei dati, alla possibile riproduzione e amplificazione di bias (pregiudizi) che creano discriminazioni, alla mancanza di trasparenza e tracciabilita', e alla vulnerabilita' ad attacchi alla sua performance.
I recenti sviluppi del deep learning hanno fornito all'IA ulteriori abilita' di generare contenuti, oltre che saperli interpretare, e soprattutto di saper gestire praticamente perfettamente il linguaggio umano, dal punto di vista sintattico. L'IA generativa permette anche la creazioni dei cosiddetti modelli fondazionali, cioe' modelli di IA molto generali che possono poi essere usati come base per costruire velocemente modelli capaci di risolvere specifici problemi. Queste abilita' espandono enormemente il panorama delle applicazioni dell'IA, con grande potenziale nell'accelerare le scoperte scientifiche e la

crescita economica, elevare il benessere della societa', e risolvere problemi globali di cruciale importanza, quali quelli relative al clima e alla salute. Permette anche alle macchine di interagire con noi in un modo molto piu' naturale. Come azienda che ha come clienti aziende in praticamente tutti i settori, all'IBM vediamo ogni giorno nuove applicazioni dell'IA nei processi aziendali, supportati da abilita' quali la creazione di sommari di documenti interni, la ricerca semantica di informazioni, la creazione di contenuti, e anche la creazione di codice. In tutti i settori e qualunque divisione aziendale, dal retail alla salute, all'amministrazione pubblica, al campo finanziario, alle telecomunicazioni, e oltre, l'IA e' in grado di migliorare le soluzioni gia' in atto e risolvere problemi prima non risolvibili.
Pero' espande anche le preoccupazioni menzionate prima e introduce nuovi rischi relativi alla generazione di contenuti pericolosi, la disseminazione di contenuti falsi ma plausibili, e la protezione di dati sensibili o coperti da copyright. Puo' anche avere un impatto ancora piu' significativo sul lavoro umano, sull'educazione, sulle attivita' creative, e sulla democrazia. Inoltre, questa tecnologia cosi' potente puo' anche essere usata da attori malintenzionati o non informati in modi non adatti che possono generare comportamenti non desiderati e pericolosi.

Come prendere il meglio dall'IA e mitigare i suoi rischi?

La prima cosa da fare e' creare un ecosistema di fiducia, giustificata e informata, verso la tecnologia e i suoi usi. Questo puo' essere ottenuto progettando e adottando politiche e leggi basate sul rischio, in cui il rischio sia associato alle applicazioni dell'IA piuttosto che alla tecnologia per se, ed imporre limiti e regole ai vari attori a seconda del loro ruolo nella complessa catena del valore dell'IA, che comprende (almeno) chi colleziona i dati, che li analizza, chi li usa per il training e il testing dei modelli fondazionali, chi costruisce i modelli piu' specifici sulla base di quelli fondazionali, chi fornisce la soluzione basata su questi modelli, e chi la usa. Ad esempio, il rischio di bias (e quindi di un trattamento non equo di alcune categorie di persone rispetto ad altre) non puo' essere individuato e mitigato adeguatamente nel modello fondazionale, ma sui modelli piu' specifici, per i quali e' conosciuto lo scopo e il contesto di applicazione. Applicazioni con piu' rischio vanno sottoposte ad uno scrutinio piu' attento, e a precise regole per tutti gli attori coinvolti.
Va anche ricordato che molti settori hanno gia' leggi mirate ad evitare discriminazioni o altri scenari non accettabili, che vanno analizzate per capire come adattarle all'uso dell'IA in questi settori.

Come identificare i rischi dell'IA?

I rischi vanno identificati attraverso un approccio inclusivo e multi-stakeholder in qui tutti gli attori della societa' vanno ascoltati: chi crea IA, chi la usa per creare applicazioni specifiche, che usa le applicazioni, chi fa ricerca in IA, le organizzazioni della societa' civile, il sistema legislativo, e chi studia l'impatto delle tecnologie sulla societa'. Gli esperti di IA, da soli, non possono identificare correttamente le preoccupazioni, l'impatto, e le implicazioni di un uso pervasivo dell'IA nella nostra societa'. Cosi' come non lo possono fare coloro che si occupano di definire le regole, e in questo senso e' benvenuta e importante l'iniziativa di questa indagine conoscitiva.

Puo' sembrare un metodo che rallenta il progresso tecnologico, ma va invece inteso come un andare piu' veloci verso lo scopo ultimo, che e' di accelerare il progresso umano tramite l'uso della tecnologia.

Basta quindi affidarsi alle leggi?

Le leggi sono necessarie, ma c'e' bisogno del contributo di tutti gli attori dell'ecosistema IA, anche perche' la tecnologia evolve molto rapidamente rispetto al processo legislativo. Le aziende che creano o usano IA devono definire adeguati approcci all'etica dell'IA, che includano formazione, governance, compliance interna, valutazione dei rischi, strumenti software, e linee guida. Chi crea soluzioni basate su IA deve anche essere trasparente sulle reali capacita', limiti, e rischi. Non perche' una legge lo richiede, ma perche' e' un vantaggio competitivo e sempre piu' un imperativo richiesto dagli altri attori della societa', inclusi i clienti.
Aziende come l'IBM hanno lavorato per svariati anni per creare strumenti tecnologici, processi interni, iniziative di formazione, prodotti, soluzioni, e piattaforme, che mettono le regole di uno sviluppo ed un uso responsabile dell'IA al centro. Un esempio e' la piattaforma piu' recente che l'IBM ha creato (chiamata watsonx), che comprende tre componenti, relative ai dati, all'IA, e alla governance, tutte fondamentali per creare una soluzione di IA responsabile.

Oltre a governi e aziende, altri attori con un ruolo importante nello sviluppo e uso responsabile dell' IA includono gli standard internazionali e i processi di certificazione, che armonizzano e danno certezze su metodi e processi. E anche i singoli utenti, che devono essere informati abbastanza per evitare un uso solo passivo della tecnologia.

Il ruolo della ricerca in IA.

La ricerca in IA, informata da consultazioni multi-disciplinari, e' una componente essenziale per mitigare i rischi dell'IA. Alcuni di questi rischi sono legati ad attuali limitazioni della tecnologia, e la ricerca mira proprio a superare questi limiti. Inoltre, la ricerca puo' anche identificare nuovi metodi e tecniche per dare all'IA nuove abilita' in un modo che sia allineato ai valori umani. L'allineamento ai valori umani e' un problema aperto e molto sentito nell'ambito della ricerca in IA, per il quale, ad esempio, la combinazione di approcci di machine learning e quelli basati su logica e conoscenza esplicita sono studiati sempre di piu'.
La ricerca in IA ha bisogno di risorse, dati, potenza computazionale, e condivisione di risultati. Per questo va supportata l'innovazione open-source che democratizza l'accesso all'IA e permette ai talenti accademici di avere un ruolo nel suo sviluppo.

Cosa puo' frenare la ricerca e le aziende italiane.

La ricerca IA nelle universita' italiane e' a livelli molto alti, con risultati innovativi e di frontiera accettati nelle migliori e piu' selettive conferenze internazionali. Questo ambiente di eccellenze va supportato nella capacita' di avere le risorse necessarie per creare e

studiare innovazione nell'ambito dell'IA, che a questo punto richiede risorse ingenti e non accessibili a tutti.

Le aziende in Italia sono per lo piu' piccole o medie, e possono essere frenate nell'adozione dell'IA da incertezze legislative, ma anche e soprattutto mancanza di fiducia nella tecnologia, di risorse, e di competenze.
La preoccupazione che l'IA possa generare contenuti falsi o che possa violare regole di copyright puo' essere mitigata dall'adozione di usi informati e adeguati al contesto applicativo. Per questo e' necessaria una formazione continua che permetta di essere consapevoli dei rischi e dei modi di affrontarli. Va anche considerato il ruolo delle grandi aziende, che mostrano sempre piu' attenzione a questi possibili freni, per esempio impegnandosi a indennizzare i propri clienti nel caso in cui l'IA violi regole di protezione intellettuale.
Inoltre, le piccole aziende spesso non hanno risorse sufficienti per creare i propri modelli di IA, ma questo non deve impedire loro di creare soluzioni di IA. I modelli open-source, ormai disponibili a migliaia, sono essenziali per mitigare questo tipo di freno all'adozione.

Il ruolo della educazione e della formazione continua.

Per un approccio responsabile allo sviluppo e all'uso dell'IA, le nuove generazioni devono essere formate per avere le competenze scientifiche e tecnologiche necessarie, ma anche per capire l'impatto che l'uso di queste competenze per creare nuova tecnologia puo' avere sulla societa'. Per questo e' necessario facilitare e incentivare un approccio multi-disciplinare alla formazione, che poi va continuato anche nella ricerca e nella formazione continua dopo gli anni di studio e nell'ambito di tutta la carriera lavorativa. Al riguardo, il ruolo delle Istituzioni e' fondamentale per sviluppare un sistema formativo orientato alla multi-disciplinarietà sia nei percorsi universitari che nella ricerca accademica. Anche in questo caso, la collaborazione tra i diversi soggetti, pubblici e provati, può favorire la creazione di quelle competenze che saranno determinanti per il lavoro del futuro.