Linee guida sull'uso responsabile dell'intelligenza artificiale

1. Premessa

L'adozione dell'intelligenza artificiale (IA), in particolare dei modelli generativi, sta trasformando profondamente il mondo della ricerca, inclusa quella in ambito di policy e think tank, e più in generale il lavoro nelle organizzazioni.
Le tecnologie IA offrono opportunità significative per accelerare i processi di raccolta dati e di analisi e produzione di contenuti, ma sollevano anche questioni etiche, legali e deontologiche che richiedono un'attenta considerazione e di cui i ricercatori e il personale di staff devono acquisire consapevolezza.
In questo contesto, l'Istituto Affari Internazionali (IAI) adotta una serie di linee guida volte a garantire un uso responsabile dell'IA, in conformità con i principi delineati a livello nazionale e internazionale, come nel caso dell'AI ACT europeo e delle Living Guidelines on the Responsible Use of Generative AI in Research della Commissione europea.

2. Principi fondamentali per l'uso dell'IA

L'uso delle tecnologie di AI generativa da parte dei ricercatori e del personale di staff deve ispirarsi a una serie di principi fondamentali. Tali principi stabiliscono se e in quali condizioni sia consentito l'uso delle tecnologie di IA per specifici processi o attività che rientrino nell'ambito di azione dello IAI.

2.1 Agentività e supervisione umana

Il principio dell'agentività e della supervisione umana evidenzia che l'IA deve essere sempre utilizzata come strumento di supporto e mai come sostituto delle attività dei ricercatori e del personale di staff. I ricercatori e il personale devono mantenere il controllo e la supervisione su tutte le fasi del processo – nel caso della ricerca, in ogni fase dalla formulazione delle ipotesi alla pubblicazione dei risultati. È essenziale che i ricercatori e il personale di staff comprendano il funzionamento degli strumenti di IA che utilizzano e siano in grado di interpretare e validare criticamente gli output generati tramite il supporto dell'IA.

2.2 Integrità

Il principio dell'integrità richiede che i risultati prodotti dalle tecnologie IA siano accurati, affidabili e ottenuti attraverso metodi etici. L'uso dell'IA non deve in alcun modo compromettere questi standard. I ricercatori e il personale devono assicurarsi che le tecnologie di IA non introducano bias o errori nei dati e che i risultati ottenuti tramite il loro utilizzo siano replicabili e verificabili; devono inoltre accertarsi che l'uso delle tecnologie IA avvenga nel rispetto della privacy di tutti i soggetti interessati e delle normative vigenti sul diritto d'autore.

2.3 Trasparenza e tracciabilità

È fondamentale che i ricercatori e il personale di staff dichiarino esplicitamente l'uso che fanno di tecnologie IA nei processi di ricerca e nelle loro attività lavorative. Ciò include una documentazione dettagliata degli strumenti utilizzati, dei dati e dei prompt forniti come input, e dei processi di elaborazione attraverso cui sono stati prodotti gli

output finali. Tali informazioni possono essere veicolate tramite degli appositi Statement da accludere in calce agli output stessi, di cui si fornisce un possibile esempio al termine di questo documento.

2.4 Responsabilità

I ricercatori e il personale sono pienamente e individualmente responsabili degli output generati con l'ausilio delle tecnologie di IA. Devono quindi essere pronti a giustificare le scelte metodologiche e operative e a correggere eventuali errori e bias derivanti dall'uso delle tecnologie di IA. Tale responsabilità implica anche la necessità da parte di ciascun ricercatore e membro del personale di staff di aggiornarsi continuamente sulle evoluzioni tecnologiche e normative relative all'IA.

---

3. Policy e raccomandazioni

3.1 Politiche e procedure

Tutti i ricercatori e il personale dello IAI sono chiamati a rispettare i principi fondamentali individuati nella sezione 2 nelle loro attività di ricerca e operative. Devono inoltre accertarsi che l'eventuale uso di tecnologie IA rispetti appieno i regolamenti, le policy e le indicazioni degli enti finanziatori e partner dei vari progetti in cui se ne prevede l'impiego, oltre che con la normativa vigente in materia a livello nazionale e comunitario.
In caso di dubbi relativamente all'effettiva rispondenza di determinati strumenti di IA o modalità di utilizzo degli stessi ai suddetti criteri, i ricercatori e il personale sono chiamati a procedere a una verifica preliminare a riguardo con il DPO dell'Istituto.

3.2 Formazione e sensibilizzazione

Lo IAI si impegna a promuovere programmi di formazione per garantire che i ricercatori e il personale di staff possano essere aggiornati sulle potenzialità e i limiti dell'IA, seguendo una linea di attività già iniziata a partire dal 2023.

---

4. Riferimenti normativi: l'AI Act dell'UE

Il Regolamento (UE) 2024/1689, noto come AI Act, rappresenta il primo quadro normativo completo sull'IA a livello europeo. L'AI Act stabilisce requisiti specifici per l'uso dell'IA, con particolare attenzione ai sistemi ad alto rischio. I ricercatori sono invitati a familiarizzare con le disposizioni dell'AI Act e assicurarsi che le loro pratiche siano conformi alle normative vigenti.
Per un approfondimento sugli aspetti legali dell'AI Act, si rimanda all'allegato di approfondimento: "**Breve guida all'uso dell'IA conforme al Regolamento UE 2024/1689**", che include una panoramica dei requisiti per i sistemi ad alto rischio, le implicazioni per la gestione dei dati e le possibili sanzioni in ipotesi di violazione.

Nota: La stesura del presente documento è stata supportata dall'utilizzo di uno strumento di intelligenza artificiale generativa (ChatGPT, sviluppato da OpenAI), impiegato sotto la piena responsabilità e supervisione dell'autore. Tutti i contenuti sono stati controllati, validati e adattati in coerenza con i principi etici e metodologici dell'Istituto Affari Internazionali.