DOI: 10.3280/for2024-003oa19285

Riconoscere e superare i bias con l'Intelligenza Artificiale

Percorsi formativi per le iniziative di Diversity, Equity & Inclusion

Marco Cocciatelli

Contributo vincitore della call for papers indetta nell'ambito di #OpenAIF in occasione del XXXVI Convegno nazionale AIF.

Introduzione

L'adozione dell'Intelligenza Artificiale (AI) nei processi decisionali e organizzativi porta con sé vantaggi in termini di efficienza, ma allo stesso tempo incrementa l'esposizione delle persone a nuovi rischi, spesso latenti. L'integrazione tra capacità umane e algoritmiche inaugura un'era di "co-intelligenza", in cui le macchine agiscono come collaboratori e coach, trasformando radicalmente il modo in cui pensiamo e lavoriamo (Mollick, 2024).

Un doppio fronte d'azione

L'interazione tra esseri umani e sistemi di AI può diffondere pregiudizi e stereotipi, influenzando negativamente le decisioni strategiche e relazionali. Algoritmi di ricerca e altre tecnologie digitali spesso riflettono le disuguaglianze sociali esistenti, amplificando stereotipi razziali e di genere (Noble, 2018). L'inserimento dell'AI nei processi aziendali comporta una trasformazione della natura del lavoro; è fondamentale che i professionisti siano in grado di riconoscere e gestire queste criticità (Daugherty e Wilson, 2018).

Una proposta metodologica integrata

Il modello formativo qui proposto si propone di fornire ai discenti gli strumenti cognitivi e operativi per riconoscere i bias e gestire efficacemente l'interazione con l'AI generativa, arginando i possibili esiti negativi dovuti alla creazione e diffusione di biased output. Il percorso si articola in tre fasi:

1. **Approfondimento teorico**: introduzione dei concetti fondamentali di bias umani e algoritmici. Attraverso case study e discussioni interattive, vengono analizzati gli impatti negativi dei bias a livello personale, relazionale e organizzativo. Gli strumenti teorici offerti creano la base per comprendere le dinamiche sottostanti e sviluppare una consapevolezza critica.
2. **Laboratori pratici esperienziali**: immergendosi in sfide pratiche basate sul roleplay, i partecipanti interagiscono con chatbot basati su AI generativa. L'esperienza diretta permette di simulare l'impatto delle distorsioni in situazioni comuni di discriminazione o pregiudizio, e di esplorare soluzioni efficaci per arginarne gli effetti negativi. Il coinvolgimento attivo consente di maturare una comprensione più profonda e critica delle tematiche e di favorire un apprendimento trasformativo (Mezirow, 2003).
3. **Sessioni di coaching e mentoring guidate da esperti**: i partecipanti ricevono feedback personalizzati e sviluppano piani d'azione per implementare le conoscenze acquisite nelle loro organizzazioni. L'obiettivo è tradurre la formazione in cambiamenti concreti, con particolare attenzione alle sfide legate alla Diversity, Equity and Inclusion (DEI).

Conclusioni

Per rispondere alle sfide poste dalla diffusione dei sistemi AI è fondamentale investire in percorsi di formazione che sappiano cogliere la complessità delle implicazioni della relazione umano-macchina.
Il coinvolgimento dei discenti in un percorso laboratoriale mediato da esperti consente di sviluppare strumenti critici e operativi per migliorare le competenze personali, relazionali e operative.

Bibliografia

Daugherty P. R., Wilson H. J., *Human + Machine: Reimagining Work in the Age of AI*, Harvard Business Review Press, 2018.

Mezirow J., *Apprendimento e trasformazione: Il significato dell'esperienza e il valore della riflessione nell'apprendimento degli adulti*, Raffaello Cortina Editore, 2003.

Mollick E., *Co-Intelligence: Living and Working with AI*, WH Allen, 2024.

Noble S. U., *Algorithms of Oppression: How Search Engines Reinforce Racism*, NYU Press, 2018.

Marco Cocciatelli
Learning content manager & designer.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see:
http://creativecommons.org