

# **Conoscere e mitigare i rischi sociotecnici dell'uso di sistemi di IA nelle politiche sociali**

*Dario Chianese (Università degli Studi di Napoli Federico II)*

Il recente interesse nell'utilizzo di sistemi di intelligenza artificiale (IA) da parte di attori pubblici e privati va commisurato ad un'adeguata concettualizzazione delle diverse implicazioni sociotecniche derivanti dal loro impiego, così da poter concertare opportune operazioni di mitigazione. Incarnando le dinamiche sociali da cui emergono, questi sistemi ne proiettano le disuguaglianze in precisi ambiti di decisione. Il loro utilizzo può arrecare danno a diversi livelli, che vanno da problemi di rappresentazione delle categorie sociali coinvolte ad effetti ambientali, politici e culturali. Particolarmente suscettibili sono l'IA generativa ed i Large Language Models (LLM), che hanno un ruolo preminente nelle implementazioni odierne. Nel discutere di questi rischi è possibile adottare come quadro di riferimento la tassonomia dei sistemi di IA sviluppata dall'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE), ed in particolare la dimensione "People and the Planet": questa riguarda in particolare gli utenti e gli stakeholder coinvolti, direttamente o indirettamente, dall'utilizzo di sistemi di IA, e suggerisce misure di disegualianza per valutare la loro esposizione ai rischi. Mobilitando questa classificazione e simili categorie di rischio sociotecnico, nonché le direttive europee ed internazionali sull'utilizzo etico dei sistemi di IA come i requisiti europei per una IA affidabile, si evidenzia concretamente come questi possano influenzare le disuguaglianze sociali e suggerire strategie per un impiego più accorto ed efficace in ambito di politiche sociali. L'utilizzo di sistemi di IA consente, infatti, di agevolare i carichi di lavoro e personalizzare il trattamento dei soggetti, purché la loro implementazione tenga conto di procedure di mitigazione che attraversano l'intero ciclo di vita dei sistemi, dalla pianificazione fino al dispiegamento ed al monitoraggio, e riguardano tutti gli attori e stakeholder coinvolti, finanche i sistemi stessi e le loro parti. Poiché i rischi considerati sono numerosi e di natura diversa è raramente possibile effettuare misurazioni dirette e le operazioni di valutazione e mitigazione possono giovare di approcci multidimensionali e transdisciplinari. In particolare, il coinvolgimento degli attori nell'implementazione e la coprogettazione possono contribuire grandemente alla messa in sicurezza dei sistemi rispetto ai rischi sociotecnici, poiché facilitano al contempo la produzione dei sistemi e la familiarizzazione di operatori ed utenti finali con essi. Pratiche di futurecasting inclusive nei confronti degli stakeholder possono rilevare criticità altrimenti difficili da intercettare o lontane nel tempo, e fanno anch'esse da volano per la produzione ed il trasferimento di conoscenze fra gli attori. Una fase di test dei sistemi e l'utilizzo di sandbox regolative offre poi l'occasione di seguire l'operazione dei sistemi di IA in ambienti controllati, effettuare valutazioni di rischio più realistiche e adattare i sistemi al contesto operativo grazie al contributo degli attori coinvolti. Attraverso una comprensione approfondita dei rischi dei sistemi di IA e l'adozione di misure appropriate è possibile implementare soluzioni tecniche trasparenti ed accessibili e chiarire gli ambiti di responsabilità e privacy, al servizio di politiche sociali più affidabili, eque e inclusive. Una focalizzazione sui rischi e le azioni di mitigazione riguardanti utenti e

stakeholder risulta perciò utile al riconoscimento e l'adozione di buone pratiche in linea con gli obiettivi strategici degli attori istituzionali.