

L'INTELLIGENZA ARTIFICIALE IN STUDIO

i vizi delle fonti non garantiscono i risultati del sistema

Martina Calcaterra, Giuliano Iannaccone e Oreste Pollicino

L'utilizzo dell'intelligenza artificiale da parte degli avvocati, oltre a presentare problematiche di accuratezza dei risultati già discusse nei contributi precedenti, solleva anche una serie di temi legati alla deontologia forense. In questo senso, risultano centrali le questioni relative al dovere di segretezza e riservatezza, al rapporto di fiducia che il professionista instaura con la parte assistita, al dovere di competenza e a quello di aggiornamento professionale e formazione continua. In particolare, in base ai doveri di competenza e formazione, è necessario per l'avvocato curare la propria preparazione relativa al tema dell'intelligenza artificiale, specie alla luce della centralità che queste tecnologie si prospettano di ricoprire all'interno della professione forense. E tale preparazione non può che comprendere, oltre a un training volto a capire il funzionamento della macchina, anche lo studio e l'attenta valutazione dei rischi legati all'uso della stessa nel lavoro quotidiano del professionista.

Uno dei rischi che più preoccupa i potenziali utenti, nonché in generale gli esperti del settore, è quello legato a eventuali bias, o pregiudizi, intrinseci alle tecnologie di intelligenza artificiale. In particolare, esistono numerose perplessità relative alle modalità di addestramento di tali tecnologie, specie se utilizzate per svolgere attività non meramente meccaniche. Come già si notava in precedenza, infatti, al fine di generare output affidabili è essenziale che il dato su cui si basa la macchina sia genuino: in altre parole, quando la fonte è precisa e imparziale, il risultato necessariamente aumenta di credibilità. Viceversa, una fonte affetta da bias e dunque per definizione parziale, se non individuata e valutata correttamente, rischia non solo di fornire output potenzialmente fallaci ma anche di perpetuare ingiustizie e discriminazioni già esistenti.

Nell'ambito delle tecnologie di intelligenza artificiale, i bias si riferiscono al trasferimento, sia esso volontario o involontario, di bias umani all'interno dei dati di addestramento forniti alla macchina. Quest'ultima, dunque, finisce per basarsi su presupposti errati e genera di conseguenza risultati distorti, che risultano difficili da riconoscere e rettificare in quanto tali distorsioni possono derivare da molteplici fattori, tra cui si annoverano disuguaglianze sociali, variabili ignorate come genere o razza e persino decisioni o comportamenti più o meno volontari dell'utente e della macchina. Si pensi, ad esempio, alla ingente quantità di dati disponibili al pubblico con cui le tecnologie di intelligenza artificiale devono interfacciarsi e a come,

talvolta, risulti più agevole persino per la macchina generare output basandosi semplicemente sulle informazioni più facilmente disponibili e più comuni all'interno del database.

© RIPRODUZIONE RISERVATA