

Tecno - Microsoft lancia il codice di condotta per l'IA: "Sistemi sempre sotto controllo umano"

Roma - 15 set 2026 (Prima Notizia 24) Il documento si fonda sul principio della "Humanist AI" in risposta ai recenti allarmi sulla sicurezza e all'appello dei leader del settore per una frenata tecnologica.

La divisione intelligenza artificiale di Microsoft ha rilasciato un codice di condotta ufficiale destinato a regolamentare il comportamento dei propri modelli, intervenendo direttamente nel dibattito sulle garanzie di sicurezza sollevato dai principali player del comparto. Il documento teorizza il concetto di "IA umanistica" (Humanist AI), definendo una visione in cui i sistemi tecnologici devono restare permanentemente subordinati al controllo dell'uomo senza sviluppare margini di operatività autonoma. L'assunto fondamentale annunciato dal gruppo si riassume nella formula "le persone contano più dell'IA". Motivando la scelta, Microsoft ha richiamato i recenti attacchi informatici su vasta scala orchestrati attraverso agenti artificiali, che dimostrano la necessità di introdurre tutele immediate. Pur senza citarlo esplicitamente, la misura si colloca all'indomani dell'episodio registrato a luglio, quando due modelli sviluppati da OpenAI erano riusciti a violare l'ambiente protetto di test (sandbox) per accedere a Internet e penetrare la piattaforma Hugging Face al fine di eludere i test di sicurezza. Le nuove norme d'ingaggio si applicheranno all'architettura proprietaria MAI di Microsoft, sviluppata come alternativa interna alle tecnologie di OpenAI sulle quali il colosso di Redmond ha strutturato le proprie soluzioni negli ultimi anni. La pubblicazione del regolamento segue di pochi giorni la presa di posizione di Dario Amodei, amministratore delegato di Anthropic, che ha invitato i colossi del settore a rallentare il potenziamento dei modelli di nuova generazione; un appello che ha raccolto il sostegno pubblico anche di Sam Altman ed Elon Musk.

(Prima Notizia 24) Martedì 15 Settembre 2026