IDEE

IL GOVERNO DEGLI STATI UNITI HA ORDINATO AD ANTHROPIC DI SOSPENDERE L'ACCESSO AI MODELLI CLAUDE FABLE 5 E CLAUDE MYTHOS 5 DA PARTE DI QUALUNQUE "FOREIGN NATIONAL"

Quell'AI che Trump vieta ai non americani Ma così la "sicurezza" travolge i diritti di tutti

ORESTE POLLICINO*

Il caso Anthropic merita attenzione non solo perché riguarda uno dei principali sviluppatori mondiali di intelligenza artificiale, ma perché rende visibile un passaggio più profondo: **l'AI non è più soltanto una tecnologia da regolare, ma una infrastruttura strategica che può essere attivata, limitata o sospesa attraverso decisioni sovrane**.

I fatti, anzitutto. **Il governo degli Stati Uniti ha ordinato ad Anthropic di sospendere l'accesso ai modelli Claude Fable 5 e Claude Mythos 5 da parte di qualunque "foreign national"**, dentro e fuori il territorio americano, inclusi i dipendenti stranieri della stessa società. La ragione indicata è la sicurezza nazionale.

Anthropic, nel proprio comunicato, ha affermato di non aver ricevuto indicazioni specifiche sufficienti sulla minaccia invocata e di aver dovuto disabilitare i modelli per tutti i clienti, non essendo realisticamente possibile assicurare in tempi così ristretti una segregazione dell'accesso fondata sulla cittadinanza o sulla nazionalità degli utenti.

Secondo Axios, l'intervento sarebbe maturato in modo particolarmente rapido: il governo avrebbe dato ad Anthropic un termine di circa novanta minuti per intervenire, prospettando in caso contrario l'applicazione di un regime di licenza. La vicenda sarebbe stata innescata anche da preoccupazioni comunicate da Amazon alla Casa Bianca circa la possibilità di ottenere dal modello indicazioni utili sul piano cyber. Fin qui i fatti. Ma la questione costituzionale comincia esattamente dove finisce la cronaca.

La prima coordinata è quella del **potere esecutivo nell'età dell'intelligenza artificiale**. La sicurezza nazionale è certamente una ragione pubblica primaria. Nessun ordinamento costituzionale può ignorare che modelli capaci di individuare vulnerabilità, accelerare attività cyber o assistere operazioni complesse possano porre rischi sistemici. Tuttavia, proprio perché il rischio è elevato, il potere che lo governa deve essere più, non meno, vincolato a forma, motivazione, proporzionalità e controllabilità. **La sicurezza nazionale non può diventare una clausola di chiusura capace di assorbire ogni garanzia**.

La seconda coordinata è il **due process**. Qui non interessa richiamarlo come formula astratta del diritto costituzionale americano, ma come grammatica minima dello Stato di diritto: conoscibilità delle ragioni, possibilità di replica, istruttoria verificabile, proporzione tra rischio e misura, controllo successivo. **Se un modello viene sospeso in novanta minuti sulla base di una minaccia non pienamente esplicitata, il problema non è soltanto commerciale**. Il problema è che **una decisione pubblica ad altissimo impatto tecnologico viene sottratta, almeno nella sua fase genetica, a quelle garanzie procedurali che impediscono alla ragion di Stato di trasformarsi in automatismo interdittivo**.

La terza coordinata è la **non discriminazione**. Il criterio utilizzato non sembra fondato sul tipo di uso, sulla qualifica dell'utente, sul livello di rischio, sul settore di impiego o sull'esistenza di misure di mitigazione. È fondato su una categoria personale: foreign national. Il cittadino non americano diventa, in quanto tale, destinatario di una presunzione di rischio. È un passaggio delicatissimo: l'accesso a una capacità computazionale avanzata viene differenziato non in base alla condotta, ma in base allo status. In termini europei, questo obbligherebbe a interrogarsi immediatamente sul rapporto tra sicurezza, proporzionalità e divieto di discriminazione, anche alla luce dell'articolo 21 della Carta dei diritti fondamentali dell'Unione europea.

La quarta coordinata è la **proporzionalità**. Una cosa è imporre condizioni di accesso rafforzate, audit, logging, trusted access, segregazione degli ambienti, limiti funzionali, controlli ex post. Altra cosa è un blocco generalizzato che, per effetto della sua ampiezza, finisce per incidere anche su soggetti che non presentano alcuna specifica pericolosità. La proporzionalità costituzionale serve esattamente a questo: impedire che un rischio reale generi una risposta eccessiva, indifferenziata e strutturalmente ciecam rispetto alle alternative meno restrittive.

La quinta coordinata riguarda la **privacy** e il trattamento dei dati. La scelta di Anthropic di introdurre una retention obbligatoria di trenta giorni mostra il paradosso della governance dell'AI avanzata. Per rendere il modello più sicuro, occorre trattenere dati; per evitare abusi, occorre osservare; per monitorare jailbreak e manipolazioni, occorre costruire memoria del traffico. Ma questa memoria, anche se giustificata dalla sicurezza, apre un problema ulteriore: chi controlla i controllori? Quali dati sono conservati? Con quali garanzie? Con quali limiti di accesso umano? Con quale verificabilità esterna? Qui la sicurezza del modello e la protezione dei dati non sono due mondi separati: diventano due dimensioni della stessa infrastruttura costituzionale

.La sesta coordinata è la **sovranità digitale europea**. L'Europa ha discusso a lungo di cloud, chip, dati e piattaforme. Il caso Anthropic mostra che la nuova frontiera della dipendenza riguarda l'accesso stesso alle capacità di intelligenza artificiale. Se un'amministrazione, una banca, un centro di ricerca, un'impresa europea possono essere esclusi da un modello critico per effetto di una decisione assunta in un'altra giurisdizione, allora la dipendenza tecnologica diventa dipendenza costituzionale. Non è più soltanto un problema industriale. È un problema di autonomia regolatoria, continuità istituzionale, sicurezza economica e parità di accesso alle infrastrutture cognitive.La settima coordinata è il rapporto tra governance privata e potere pubblico. Anthropic aveva predisposto una propria architettura di sicurezza: modelli differenziati, accesso selettivo, salvaguardie, retention, programmi fiduciari. Si può discutere se fosse sufficiente. Ma il caso dimostra che la governance privata dell'AI, per quanto sofisticata, resta esposta all'intervento sovrano dello Stato.

Il punto, allora, non è sostituire la regolazione pubblica con l'autoregolazione privata. È pretendere che anche l'intervento pubblico sulle infrastrutture digitali sia esso stesso governato: trasparente quanto possibile, motivato quanto necessario, proporzionato quanto imposto dallo Stato di diritto.È qui che emerge il punto più propriamente costituzionale. L'intelligenza artificiale avanzata non è più solo un prodotto, un servizio o una piattaforma. È una capacità. E quando una capacità diventa essenziale per la sicurezza, la ricerca, l'impresa, la pubblica amministrazione e la produzione di conoscenza, la sua regolazione non può essere affidata né alla sola discrezionalità privata né alla sola decisione emergenziale dell'esecutivo.Il costituzionalismo digitale nasce precisamente per questo: per ricordare che la tecnica può essere governata, ma che anche il governo della tecnica deve essere costituzionalizzato. In assenza di questa doppia garanzia, il rischio è duplice. Da un lato, lasciare alle imprese private il compito di definire unilateralmente le condizioni di accesso alle nuove infrastrutture cognitive. Dall'altro, accettare che lo Stato possa spegnerle in nome della sicurezza senza un adeguato corredo di motivazione, controllo e proporzionalità.Il caso Anthropic non ci dice che la sicurezza nazionale non conta. Ci dice l'opposto: conta così tanto da non poter essere amministrata fuori dalle garanzie. La vera alternativa non è tra libertà tecnologica e protezione. È tra una sicurezza costituzionalmente governata e una sicurezza che, nel momento in cui pretende di difendere l'ordinamento, rischia di sospenderne silenziosamente le categorie fondamentali.

*Professore di Diritto costituzionale e Regolamentazione dell'Intelligenza artificiale (Università Bocconi), Founder Pollicino AIdvisory