I SEGRETI DELL'INTELLIGENZA ARTIFICIALE: UNO SCONTRO DI POTERI

Le implicazioni sociali, economiche e politiche dell'IA. I rischi associati alla sua presenza nascosta e pervasiva nella quotidianità e all'esercizio del controllo. L'algoritmo come segreto industriale e pubblico e la questione dell'eticità.

Claudio Franchi

Due tra i profili più problematici sull'Intelligenza Artificiale sono correlati con sfaccettature diverse al dispositivo del segreto: la segretezza assoluta dell'algoritmo che la fa funzionare e la dimensione occulta della sua pervasività che addirittura impedisce anche la semplice percezione della sua esistenza e operatività. A questi due nodi si aggiunga poi un elemento incontrollabile e imprevedibile ma presente e ineliminabile: le reti neurali, essenziali per il funzionamento dell'IA, molto spesso hanno un "luogo", un "momento", che non è assolutamente incluso nelle programmazioni scritte dagli umani, dove fanno valutazioni statistiche e prendono decisioni rispetto ai propri output, che possono essere anche delle vere e proprie azioni. Anche secondo i programmatori, in questo caso, ci troviamo di fronte a una vera e propria scatola nera, una *Black box*, della quale ignoriamo gli esiti e lo stesso improvviso apparire.

Una microgenealogia del potere politico del segreto

Proviamo però a comprendere quale può essere una lente che ci permetta di capire con maggiore efficacia quali siano le poste in gioco e in che ambiti ci si muova. Una tecnologia come l'IA ha troppe conseguenze e implicazioni sociali, economiche e soprattutto politiche. Tentiamo quindi di utilizzare qualche chiave interpretativa desunta dalla teoria politica. Se solo si volesse tracciare una piccola genealogia almeno delle pietre miliari del "segreto" in questo ambito, ci troveremmo di fronte innanzitutto agli *Arcana Imperii* di Tacito, poi velocemente passeremmo per Machiavelli e già arrivati a Hobbes, capiremmo che si sta trattando un tema ritenuto e posto come fondamentale e che i suoi meccanismi sono essenzialmente una modalità quasi strutturale per l'esercizio del potere, soprattutto di chi quel potere lo possiede già e lo esercita.

È proprio questo il punto chiave. La lettura necessaria dell'Intelligenza artificiale infatti, al di là delle caratteristiche tecniche e tecnologiche e delle eventuali modalità di applicazione, deve vederla come un vero e proprio spazio nel quale si intersecano dinamiche politiche, economiche e sociali, istanze democratiche e anti democratiche, forme di liberazione dal lavoro e procedure esasperate di controllo e autoritarismo, interessi meramente economici e progressi reali della conoscenza, per non elencare che alcuni vettori principali nella fisica dei poteri in gioco. Il segreto è esattamente uno dei meccanismi di questi vettori.

Per Machiavelli, nel *Principe*, il segreto è uno strumento strategico per il mantenimento del potere e i governanti, secondo le sue parole, dovrebbero "talvolta" - cioè tutte le volte che è necessario - ricorrere a tattiche segrete per raggiungere i loro obiettivi politici. Quando Thomas Hobbes invece, nel *Leviatano*, discute del contratto sociale e del potere sovrano, il segreto è spesso associato alla necessità di mantenere l'ordine sociale e nel momento in cui sostiene che il sovrano debba sapere utilizzare il segreto per preservare la stabilità del

governo, dal nostro punto di vista la coincidenza di governo e ordine sociale - basati entrambi sul segreto - delineano ancora di più le coordinate da utilizzare per analizzare con chiarezza e consequenzialità questo fenomeno rispetto all'IA.

Quando allora iniziamo a includere nella nostra breve disamina autori più recenti, Foucault, in *Sorvegliare e punire*, va ancora più a fondo e esplorando le dinamiche del potere e del controllo sociale, evidenzia come i segreti siano addirittura parte *integrante* delle relazioni di potere nella società: «la società deve essere analizzata come un complesso di relazioni in cui il segreto è una delle forze attive». Giorgio Agamben, invece, sposta il suo sguardo dalla normale vita sociale e politica al concetto di "stato d'eccezione", indagando il legame tra il segreto e la gestione dello Stato in situazioni di emergenza. In questo caso, il segreto può essere utilizzato nell'attuazione e al contempo nella giustificazione di misure straordinarie - cioè potenzialmente fuori dai limiti della legge, nel contesto politico dato.

Gli sguardi, gli scopi, dei pensieri presi in considerazione possono essere diversi rispetto ai ruoli e funzioni possibili del segreto in relazione alla dimensione politica dei poteri, oppure rispetto a come esso possa essere utilizzato di volta in volta per il controllo, la stabilità sociale o la giustificazione di azioni "straordinarie". Ma non potremmo mai esimerci dal citare le parole di Hannah Arendt ne *La banalità del male:* «il segreto è la radice di tutto il male». Ovviamente la Arendt sta esaminando la natura della cattiveria umana, ma è proprio qui che sottolinea come i segreti possano contribuire alla perpetuazione di azioni totalmente contrarie all'etica, qualsiasi forma esse prendano e in qualsiasi veste ideologica siano proposti nel discorso pubblico.

L'algoritmo come segreto industriale e pubblico

Il segreto industriale, per definizione, è un'informazione che conferisce un vantaggio competitivo a un'impresa. Questo meccanismo, centrale nelle economie capitaliste e proprio per questo protetto attraverso leggi che lo difendono, permette di impedire la diffusione o la sottrazione di informazioni critiche su processi di produzione, formule chimiche, design o addirittura strategie di marketing, tra le altre cose. Se si dovesse citare, un esempio tra gli altri, un'azienda che produce una storica bevanda diffusa in ogni paese del mondo, ne protegge la formula chimica o il processo di produzione per impedire a eventuali concorrenti di copiarlo. Questo consente all'azienda di mantenere un reale vantaggio economico sul mercato, mantenendo esclusività e preservando una propria posizione dominante, difesa dai quadri legislativi nazionali e internazionali.

Un interrogativo possibile però dovrebbe nascere spontaneo: questo segreto è eticamente corretto? La questione dell'eticità del mantenimento dei segreti industriali dipende ovviamente molto dal contesto e dalle circostanze specifiche, oltre che, ovviamente, dalla tipologia di sistema economico nel quale nasce. In generale, mantenere i segreti industriali è considerato etico finché non si violano leggi o accordi contrattuali. Tuttavia, ci possono essere - e ci sono stati - casi in cui la divulgazione di informazioni riservate potrebbe essere giustificata da motivi di interesse nazionale, come ad esempio per proteggere la sicurezza dello Stato o per promuovere e difendere la salute pubblica.

Un caso ipotetico potrebbe essere uno Stato che ha bisogno di informazioni su una nuova tecnologia farmaceutica segreta sviluppata da un'azienda nel suo territorio per combattere un'epidemia grave. In questo caso, potrebbe essere considerato etico, per lo Stato, richiedere la divulgazione delle informazioni al fine di produrre più rapidamente un trattamento efficace per proteggere la salute pubblica.

Assistiamo qui chiaramente, guardando al di là del dato regolativo, a un fortissimo conflitto fondato esclusivamente su interessi economici che, in un quadro anche democratico di bilanciamento di poteri diversi, viene considerato di default come spostato dalla parte delle imprese economiche, salvo situazioni o di catastrofe possibile o di necessità di mantenere lo *status quo* politico, come tra l'altro abbiamo visto sopra nella nostra breve microgenealogia. Da un punto di vista preciso, viene ritenuto importante "bilanciare" qualsiasi tipo di interessi, anche ipoteticamente "pubblici" con il rispetto assoluto per i diritti di proprietà intellettuale e i diritti commerciali dell'azienda.

È questo il quadro nel quale si staglia la questione degli algoritmi proprietari: tutte le aziende considerano gli algoritmi di intelligenza artificiale come segreti commerciali e ciò viene posto come ostacolo insormontabile e protetto dalle leggi, permettendo di sorvolare amabilmente sulle preoccupazioni circa la trasparenza e l'equità, a prescindere dall'impatto che le decisioni dell'IA possono avere sulla vita delle persone senza che nessuno possa comprendere appieno come esse vengono prese.

La questione diviene però ancora più complessa e soprattutto pericolosa, quando è lo stesso Stato che utilizza questi algoritmi per analizzare, ipoteticamente, masse enormi di dati e informazioni e prendere poi decisioni che investono e modificano la vita di cittadine e cittadini, rifiutandosi in ogni modo di rivelare quali siano stati i parametri che hanno portato poi a queste stesse decisioni, sotto il paravento ideologico della "attribuzione di responsabilità" a i dati presentati come oggettivi. Se si guardasse solo alle questioni italiane, nel 2016 un algoritmo "segreto" avrebbe deciso la mobilità nazionale - tecnicamente, il trasferimento interprovinciale - dei docenti, sconvolgendo vite personali, affetti e economie materiali, senza nessun confronto con i rappresentanti sindacali, laddove il Governo si è trincerato dietro le decisioni "automatiche" dell'algoritmo, che invece poteva e doveva essere condiviso con le parti sociali. E anche le sentenze del TAR relative hanno semplicemente assicurato un accesso agli atti *ex post*, come visione e estrazione di copia, ma lasciando completamente nelle mani dei decisori politici l'intero processo di elaborazione e implementazione dell'algoritmo. Una situazione analoga si è verificata durante la pandemia, quando si dovevano stabilire le forme di restrizione alle libertà su base regionale per cercare di evitare ulteriori contagi: il governo in carica allora decise autonomamente e senza nessun confronto reale quali fossero i parametri da inserire nell'algoritmo e come questo dovesse funzionare. Agisce in questo caso lo stesso meccanismo che nel medioevo era il principio di *auctoritas*, solo che al posto di Aristotele, che almeno poteva essere accessibile, ci sono oggi gli algoritmi e il principio solo ideologico di "oggettività scientifica" difesa dal segreto, e che in realtà contravviene esattamente a uno dei principi cardine della scienza che è quella della condivisione pubblica di dati e processi

per permettere la riproducibilità degli esperimenti e quindi l'effettiva verificabilità. Altrimenti si deve parlare di Fede - o, meglio, di *Malafede* - e non di Scienza.

Nota: non è possibile in questo luogo parlare né della *privacy* dei dati che gli algoritmi di intelligenza artificiale usano per "allenarsi" e per produrre output - la questione della *privacy* sorge quando queste informazioni vengono utilizzate per addestrare modelli, specialmente se i dettagli personali sono coinvolti -, né dei possibili *bias* impliciti nei modelli, non controllabili a causa del segreto nella progettazione degli algoritmi, che riflettono e perpetuano pregiudizi presenti nei dati di addestramento, determinando discriminazioni ingiuste o amplificando disuguaglianze esistenti.

L'Intelligenza Artificiale come presenza pervasiva e nascosta

Scriviamolo chiaramente, la possibile pervasività del "controllo" - inteso come esito nelle relazioni umane della raccolta sistematica di dati e informazioni - operato dai sistemi di automatizzazione algoritmica e di IA incrementa in modo esponenziale la capacità sempre maggiore attraverso l'uso cosciente di tali sistemi di influenzare e governare varie sfere della vita umana, senza che si abbia la benché minima percezione del processo in atto da parte dei singoli individui. E questo a prescindere da eventuali *disclaimer* ai quali possa essere stata data visione o adesione, in quanto nella pratica quotidiana essi non sono un reale avvertimento, ma piuttosto una salvaguardia legalistico-formale delle imprese, per rendere inefficaci le leggi e difendersi dalle possibili azioni legali o risarcitorie.

Una delle possibili manifestazioni di questo fenomeno, che coinvolge un alto grado di automazione, sono gli algoritmi di raccomandazione - acquisti, video, film, canzoni, contatti sui social - che influenzano le nostre scelte proponendo elementi che possano essere di nostro gradimento e producono questi "suggerimenti", surrettiziamente proposti come "personalizzazione" o miglioramento delle esperienze. Questi però possono essere generati solo grazie al controllo completo e all'analisi immediata - possibile proprio grazie all'enorme accumulo e integrazione di dati e di potenza di calcolo dei sistemi automatici - delle nostre attività, dai video già visti, all'utilizzo delle carte di credito, dall'identificazione e la quantificazione dei secondi passati a guardare un singolo post sui diversi social, al controllo incrociato tra le nostre rubriche, i nostri contatti e le nostre attività. Allo stesso modo funzionano i sistemi di sorveglianza che monitorano le nostre presenze fisiche, o decisioni automatizzate che possono avere impatti significativi sulla nostra vita quotidiana.

I rischi associati a questa pervasività, molto spesso occultata, vengono di solito ricondotti a una massiccia perdita di *privacy*, che in realtà viene però quasi esplicitamente barattata con l'utilizzo dei servizi diversi, a partire dall'utilizzo dei social per finire alle app di geolocalizzazione o di navigazione, tutti ufficialmente gratuiti ma invece "pagati" attraverso la messa a disposizione dei propri dati personali e di comportamento, che costituiscono una vera e propria moneta sonante nel mondo delle tecnologie. Quello che invece accade veramente - e che nel discorso pubblico viene spesso sottovalutato o accantonato - è l'accentuazione ulteriore delle disuguaglianze sociali ed economiche, che sfocia in una completa perdita di controllo democratico e effettivo su decisioni critiche, costruendo un

pericolosissimo potenziale per abusi da parte di governi o, peggio, soggetti con volontà specifiche.

Se si guarda per esempio al mondo del lavoro, un numero sempre maggiore di aziende utilizza sistemi di sorveglianza basati sull'IA, come telecamere intelligenti e software di analisi comportamentale, per monitorare i dipendenti sul luogo di lavoro. Questi sistemi possono essere utilizzati per tracciare il tempo di lavoro, valutare le prestazioni dei dipendenti e identificare comportamenti ipoteticamente non conformi. Come già sopra accennato, ufficialmente ciò solleva preoccupazioni, spesso insormontabili, riguardo alla privacy, dietro il presunto aumento della sicurezza dei dipendenti o della produttività, ma in realtà rende ancora più asimmetrico il peso del potere di controllo all'interno di un luogo o di un rapporto di lavoro. In questo contesto anche l'automazione dei processi decisionali utilizza in forma decisiva una dimensione segreta della quale i soggetti che subiscono le decisioni non sono al corrente e che influenza spesso in modo determinante le loro vite. Molte aziende o istituzioni utilizzano infatti algoritmi di intelligenza artificiale per assistere o addirittura prendere decisioni riguardanti azioni successive, come per esempio le assunzioni di nuovi dipendenti. In questo caso particolare, gli algoritmi analizzano i curriculum, i test di valutazione e altre informazioni pertinenti per identificare i candidati migliori o allo stesso tempo quelli "peggiori", cioè potenzialmente non in linea con gli interessi di parte dell'azienda, come per esempio soggetti che potrebbero essere sindacalizzati o sindacalizzabili. Quindi in questo caso, non c'è solo il rischio di discriminazione involontaria, qualora gli algoritmi siano stati addestrati su dati storici che riflettono pregiudizi di genere, razza o altro, ma proprio l'utilizzo strategicamente consapevole e volontario di un'IA per estrarre dati interpretativi, sfruttando una posizione di forza economica e sociale per azioni eticamente più che discutibili e politicamente condannabili.

Ovviamente, ci sono anche vantaggi significativi nell'uso dell'IA. Soprattutto sono notevoli i miglioramenti nell'efficienza e nella produttività, ma queste categorie devono assolutamente essere precisamente definite in quanto troppo generiche per rappresentare un vero elemento ermeneutico, e non possono essere ristrette solo a un aspetto economicista in senso stretto. Proprio se si allargano questi concetti le nuove funzioni dell'IA, sinora immaginabili solo nella narrativa della fantascienza, possono presentare nuove opportunità di innovazione e progresso scientifico. E se si volesse solo prenderle in considerazione quali potenziali soluzioni per problemi complessi in ambiti essenziali per l'umanità, come il trattamento di alcune malattie, già sarebbe un radicale miglioramento reale della vita di donne e di uomini.

La questione però non è semplicemente quella di mitigare i rischi attraverso un quadro legislativo e nello stesso tempo massimizzare i vantaggi, in virtù di un approccio quasi positivista all'evoluzione umana e sociale. Infatti da una parte è certamente essenziale sviluppare e attuare normative etiche e legali rigorose, garantire la trasparenza e l'*accountability* degli algoritmi, e anche promuovere l'istruzione e la consapevolezza pubblica sull'IA. Ma il vero snodo, soprattutto se visto con la lente prismatica del tema del segreto, è il reale coinvolgimento delle diverse soggettività sociali - in alcune parti,

addirittura chiaramente, parti sociali - che sempre hanno interessi diversi se non addirittura opposti e confliggenti, interessate nella sua progettazione e implementazione.