

L'allarme lanciato da alcune grandi società "big tech" (Anthropic, OpenAI, Google deepmind) sui seri rischi di uno sviluppo autonomo dell'Intelligenza artificiale (IA) fuori del controllo umano ha drammatizzato la discussione già ampia ed accesa in atto nel mondo su questo salto di qualità nello sviluppo dell'informatica, produttivo di massicci effetti su una quantità di questioni di grande rilievo.

La Chiesa cattolica sta manifestando di avere un particolare interesse all'argomento, con l'Enciclica di Papa Leone e la produzione di interventi, saggi, articoli da parte di personalità ecclesiastiche e laico-cattoliche. Non solo per le conseguenze dell'IA sul lavoro, la cultura, la politica, ma per il suo impatto sul funzionamento dell'intelligenza umana, componente essenziale della spiritualità individuale e collettiva la cui cura è la ragione d'essere della Chiesa. Papa Leone XIV ha sottitolato la sua prima Enciclica, "*Magnifica humanitas* ", con l'espressione "*Sulla custodia della persona umana nel tempo dell'intelligenza artificiale* ". La posizione della Chiesa, in estrema sintesi, è questa: "l'essere umano è relazione e destino, non calcolo" e occorre impedire che con l'IA "si perda l'essere umano"¹.

L'IA produce risultati positivi di grande rilievo potenziando una serie di attività umane, ma anche seri problemi per il lavoro e i diritti delle persone. In questa sede si danno per noti gli uni e gli altri, per mettere a fuoco la questione di fondo del confronto, della relazione, dell'impatto reciproco tra IA e intelligenza umana che va approfondita riflettendo in parallelo su entrambi i fenomeni, sia sul versante antropologico che su quello politico-sociale. A questa riflessione si intende qui fornire un contributo, senza pretendere di spiegare tutta la complessità dell'insieme.

L'IA è la capacità di un sistema composto da strumentazioni e metodologie tecnologiche e scientifiche di riprodurre processi mentali propri dell'intelligenza umana, come la capacità di analizzare, generalizzare, ragionare, imparare dall'esperienza, per realizzare gli obiettivi assegnatigli dal creatore - umano - del sistema. Questa riproduzione avviene attraverso il funzionamento di reti neurali analoghe a quelle del cervello umano. Queste sono composte da cellule nervose - i neuroni - che ricevono, elaborano e trasmettono informazioni attraverso segnali elettrochimici. La ricezione e la trasmissione avvengono attraverso le sinapsi, ovvero i filamenti che collegano i neuroni, componendo le reti neurali biologiche di ogni cervello umano. I neuroni di ogni cervello sono circa 100 miliardi, le sinapsi migliaia di miliardi, per cui i collegamenti possibili sono una cifra astronomica. Naturalmente, la stragrande maggioranza dei neuroni non viene mai attivata, restando sullo sfondo dell'attività cerebrale con le sue potenzialità pressoché infinite.

Le reti neurali dell'IA sono composte da unità di calcolo che non sono entità fisiche ma funzioni matematiche, collegate tra loro per ricevere, elaborare e trasmettere informazioni attraverso impulsi elettrici, analogamente a quelle del cervello umano. Oltre a questa differenza materiale c'è una differenza strutturale. Entrambi i tipi di rete si articolano su tre strati di neuroni, che corrispondono alle tre diverse funzioni di ricezione, di elaborazione e di trasmissione. In uno strato vi sono neuroni che ricevono input, ovvero elementi di informazione iniziale, in un altro quelli che elaborano questi elementi, in un terzo strato quelli che trasmettono all'esterno i risultati. I neuroni dello strato di elaborazione sono connessi sia ai neuroni dell'input sia a quelli dell'output. La differenza strutturale con

l'intelligenza umana è che questo strato intermedio nelle reti neurali dell'IA può essere affiancato da altri strati intermedi, tutti interconnessi, moltiplicando in modo esponenziale le capacità di elaborazione del sistema.

Questa moltiplicazione potrebbe condurre a forme più complesse di IA. L'Intelligenza Artificiale Generale è una forma ipotetica di intelligenza artificiale in grado di comprendere, apprendere, ragionare e applicare le conoscenze in qualsiasi ambito intellettuale allo stesso livello, o anche superiore, a quello di un essere umano. Avrebbe una capacità di generalizzazione, ossia la dote di trasferire ciò che ha imparato in un contesto a un dominio completamente diverso, adattandosi a novità e imprevisti come farebbe una intelligenza umana.

La Superintelligenza Artificiale sarebbe un ipotetico sistema informatico dotato di capacità cognitive e intellettuali nettamente superiori a quelle della mente umana, con forme di intuizione, creatività e capacità di problem solving qualitativamente inimmaginabili per il cervello biologico umano. La base dovrebbe essere l'"Automiglioramento ricorsivo" teorizzato dal filosofo Nick Bostrom², per cui, una volta raggiunta un'intelligenza pari a quella umana, l'IA sarebbe in grado di riscrivere il proprio codice e di riprogettarsi da sola a velocità esponenziale, portando a una rapida "esplosione di intelligenza".

In realtà, la differenza più importante tra le due intelligenze è un'altra. Ogni rete neurale cerebrale è aperta non solo ad altre reti neurali dello stesso cervello, come le reti artificiali, ma attraverso queste può ricevere input provenienti dalla mente profonda, ovvero da una parte degli altri neuroni dormienti, non attivi ma comunque contenenti altri elementi di informazione. Impulsi, emozioni, suggestioni, lontani ricordi che arrivano alla coscienza, alle sue elaborazioni razionali, alla memoria cosciente, dalla dimensione del subconscio e dell'inconscio. Entrando nelle reti neurali, arricchendole, modificandole. Ciò avviene non solo nelle menti individuali, ma anche a livello collettivo. Le menti individuali sono collegate tra loro non solo nella dimensione cosciente-razionale, attraverso la comunicazione orale, scritta, audiovisiva, di informazioni e concetti, ma anche nella dimensione più profonda, dall'empatia come comunicazione e condivisione immediata di emozioni ed altri contenuti dell'inconscio. Vittorio Gallese nel suo lavoro più recente³ constata che l'intelligenza umana non è una facoltà solitaria, ma prende forma nell'incontro con gli altri anche a livello profondo. Con la intersoggettività emotiva primaria, che produce la sintonizzazione tra bambino e genitore, e con la risonanza tra i "neuroni specchio", che in una persona attivano emozioni simili a quelle percepite nel contatto con un'altra persona. Dimostrando così che anche sul piano neurofisiologico opera una connessione immediata tra l'io e il tu. La dimensione razionale e quella profonda, come il livello individuale e quello collettivo, non sono staccate se non concettualmente, e dalle loro interrelazioni si producono intuizioni, creatività, premonizioni.

Le reti neurali dell'I.A., naturalmente, nel loro funzionamento ignorano le emozioni umane come ansia, timore, volontà di potenza, e neppure riproducono artificialmente comportamenti ispirati a queste emozioni. L'energia è elettrica, fornita dall'esterno, e la spinta a muovere è quella impressa all'inizio, al momento della loro creazione. Queste reti sviluppano miliardi di collegamenti orizzontali secondo algoritmi segmentati non per parole,

simboli o concetti, ma in frammenti di dati secondo frequenze statistiche di co-occorrenza di caratteri e byte. Per Raul Gabriel *“E’ il dominio del significante, liberato dalla necessità del significato”*⁴. Per Papa Leone, l’Intelligenza così assolutizzata *“finisce con l’oscurare le altre dimensioni della vita: affetto, volontà, dedizione, relazione”*⁵.

Il problema – enorme – è che l’IA cresce su sé stessa, impara da errori ed esperienze, si autonomizza rispetto all’intervento umano. L’intelligenza umana assegna al singolo sistema IA da lei creato funzioni e obiettivi per sviluppare la produzione di merci e servizi utili alla vita umana e sociale, in una logica di concorrenza rispetto ad altri sistemi di IA o comunque di produzione. Il singolo sistema così orientato può articolare gli obiettivi ricevuti in sub-obiettivi, sulla base delle esperienze realizzate, in modo automatico, matematico, assolutamente senza valutazioni etiche o valoriali. E la realizzazione di questi sub-obiettivi può condurre ad individuare target da colpire, scelti in autonomia fuori dal controllo umano, che possono essere altri sistemi di IA (è già accaduto), e potenzialmente altri esseri umani. È qui la svolta epocale, il punto in cui l’accumulazione e l’intreccio di cambiamenti quantitativi produce un radicale, negativo salto di qualità. Gli stessi padroni dell’IA, ovvero le aziende Big Tech, pongono la necessità di regolarne lo sviluppo, perché rischiano di perderne la padronanza.

Per capire come e dove intervenire è necessario analizzare i processi di cambiamento che hanno portato a questo punto, per quanto complessi possano essere. A questo riguardo può essere utile ricorrere ad alcuni strumenti concettuali del pensiero critico.

Il primo è il ciclo D (denaro)-M (merce)-D1 (denaro) dove D1 è maggiore di D, ovvero la legge fondamentale dell’accumulazione capitalista secondo Marx. Il secondo proviene da un pensiero critico più antico, l’insegnamento di Buddha, per il quale la matrice della sofferenza umana è la “sete”, l’avidità appassionata, la spinta alla soddisfazione e all’affermazione di sé. L’intreccio dei due concetti porta a considerare la spinta intrinseca all’autovalorizzazione del capitale come “sete” oggettivizzata, tendenza generale alla crescita impressa alle attività umane attraverso la “sete” soggettiva degli individui⁶. Questa spinta si esercita non solo come propulsiva dell’intero ciclo, ma investe anche ciascuno dei termini, Merce e Denaro, in sé. È la logica del profitto e dello sviluppo che ha prodotto sia la moltiplicazione delle merci, dei beni e dei servizi offerti sul mercato in cambio di denaro, sia la crescita esponenziale dei capitali, divenuta metastasi con la finanziarizzazione dell’economia e la finanza derivata, ovvero l’autonomizzazione dello sviluppo dei capitali finanziari in sé rispetto ai profitti derivanti dalla produzione di merci.

In particolare la moltiplicazione delle merci, funzionale alla crescita dei capitali, ha riguardato sia l’offerta che la domanda di beni e servizi indirizzati a soddisfare i bisogni e le esigenze derivanti dalle attività umane. Tra i servizi si sono sviluppati quelli prodotti dall’ICT, merci materiali e immateriali offerte sul mercato come strumenti per potenziare le attività di produzione delle imprese e di comunicazione delle persone, prodotti informatici spesso sostitutivi di attività umane complesse. Tutte queste direttrici e modalità di sviluppo

sono state realizzate dall'intelligenza umana, mossa dalla "sete" soggettiva dei protagonisti e dalla "sete" oggettiva dei capitali. L'IA in senso proprio, ovvero l'avvento di sistemi capaci di apprendere dalle proprie esperienze e di correggere il proprio operare, si è realizzata anch'essa per la spinta a crescere impressa all'origine in un contesto di competizione con altri sistemi, in collegamento con la spinta a crescere dei capitali finanziari investiti nell'IA, anch'essi in competizione tra loro. E le due spinte si sono sommate, potenziandosi a vicenda. Si è prodotta, così, una nuova forma di autonomia della merce rispetto al processo D-M-D, come autonomia dell'intelligenza artificiale, la merce, rispetto all'intelligenza umana alla guida dell'intero processo, che poi investe tutte le attività umane. È questo ruolo dell'intelligenza umana che viene messo in discussione, in molti settori della produzione di merci e della valorizzazione dei capitali. Col rischio, o la quasi certezza, che la competizione porti l'IA a colpire interessi delle aziende e interessi delle persone, senza intervento umano. La consapevolezza di questo rischio ha prodotto la denuncia delle Big Tech e la richiesta di una regolamentazione dello sviluppo dell'IA. Dal punto di vista di queste, ovvero dei grandi capitali impegnati nel settore, il rischio assume diversi profili. In prima linea c'è il rischio di attacchi incontrollati tra sistemi, già verificatisi in qualche caso. Sullo sfondo c'è il rischio che scoppi la gigantesca bolla finanziaria creata dai debiti fatti per moltiplicare i data center, nella prospettiva di uno sviluppo senza fine delle potenzialità dell'IA. Prospettiva messa in forse dagli attacchi incontrollati e anche dalla nascita di nuove imprese che offrono sistemi IA a prezzi inferiori a quelli dell'oligopolio big tech. E anche questo caso si è già verificato. La regolamentazione richiesta, perciò, dovrebbe investire soprattutto questi profili.

Trump, anche su questo argomento, sta dimostrando di essere l'espressione in purezza della "sete", dell'avidità, dell'affermazione esasperata di sé che aggrava le condizioni del mondo. In un primo tempo ha respinto l'ipotesi di una regolamentazione, in linea con le posizioni precedenti in materia, per favorire uno sviluppo senza limiti che assicuri il primato degli USA. Inoltre, ha annunciato che i documenti ufficiali del governo degli Stati Uniti non utilizzeranno più il termine "Artificial Intelligence" (Intelligenza Artificiale), sostituendolo ufficialmente con l'espressione "Super Intelligence" (ovvero, intelligenza superiore o superintelligenza). Non solo per esigenze di comunicazione, perché "superiore" suona meglio di "artificiale" e fa vendere di più. È una affermazione dal significato antropologico, di superiorità della IA, anche al livello attuale, rispetto all'Intelligenza umana, più lenta e meno performante perché proveniente dalla complessità della psiche.

Successivamente, dopo un confronto con le aziende big tech, le cui richieste non poteva ignorare, Trump ha recepito la necessità di una regolamentazione dello sviluppo dell'IA, secondo la sua linea e in accordo con gli interessi delle grandi imprese. Non per regolare l'innovazione e lo sviluppo tecnologico con vincoli burocratici, ma affidando la gestione dei rischi all'autodisciplina delle imprese. L'intervento dello Stato viene affidato ad un personaggio che già dirige l'Intelligence, ovvero il coordinamento delle Agenzie dei servizi segreti, e che ora dovrà anche guidare, come uno "Zar", la Super Intelligence Force, una nuova task force istituita ad hoc dalla Casa Bianca. L'affidamento al Capo dei Servizi segreti è significativo, perché una prima missione della "Force" sarà quella di proteggere le

infrastrutture e i sistemi da attacchi informatici, sabotaggi tecnologici, intrusioni e violazioni. L'altra missione, strettamente intrecciata, è quella di vincere la gara in atto con la Cina su questo campo, garantendo la superiorità americana. Questa impostazione conflittuale a livello globale, che aggrava quella concorrenziale già in atto tra le imprese, aumenta ancora i rischi di cui s'è detto, di una elaborazione autonoma di sub-obiettivi da parte dei singoli sistemi di IA indirizzati alle funzioni di attacco e difesa, a danno non solo di altri sistemi IA e delle loro funzioni, ma anche delle condizioni di vita, e della vita stessa, di parti consistenti del genere umano.

Per contro la Cina, respingendo ogni ipotesi di limitazione esterna al proprio sviluppo nel settore, postula una "governance" globale che garantisca la sicurezza e gli interessi di tutte le parti in campo, in armonia col suo indirizzo generale nelle relazioni internazionali teso a privilegiare la diplomazia multilaterale e ricercare compromessi vantaggiosi con tutti gli interlocutori.

Il Papa, nell'Enciclica già citata, chiede trasparenza nelle decisioni, individuazione delle responsabilità e una governance diffusa anche nelle comunità e dei corpi intermedi attraverso controlli ispirati ai principi di partecipazione e sussidiarietà, per contribuire al discernimento e alla vigilanza sull'addestramento dei sistemi. Soprattutto, chiede di "disarmare la IA", evitando la logica della competizione armata, militare, economica e cognitiva, e rompendo l'equivalenza tra potenza tecnica e diritto a governare, propria dei grandi monopoli dell'IA. Declinando così in indirizzi politici precisi, anche per la materia dell'IA, la visione umanista di cui s'è detto, inquadrata nella prospettiva cristiana della trascendenza.

Affrontare la rivoluzione prodotta dallo sviluppo dell'IA richiede necessariamente uno sviluppo dell'intelligenza umana, che Mons. Paglia, nella sua prospettiva, chiama "grande risveglio spirituale"⁷. Ma un indirizzo di fondo che vada in questa direzione può definirsi anche in una prospettiva più laica. A livello individuale si tratta di espandere la connessione della coscienza razionale con la mente profonda, per aumentare la consapevolezza dei contenuti di questa dimensione interiore. È quella che viene definita "intelligenza emotiva", ovvero il controllo cosciente delle emozioni e degli impulsi irrazionali⁸, decisivo per governare queste pulsioni oggi sollecitate dall'Infosfera del web, potenziata dalla IA, per indirizzarle ai fini economici e politici del tecnocapitalismo.

A livello collettivo, come s'è detto, si realizzano collegamenti con la psiche delle altre persone sia nella sfera razionale, attraverso le varie forme di comunicazione, anche interattive, sia nella sfera più profonda, attraverso l'empatia, la compassione, l'intuizione e il riconoscimento dell'altro. Anche qui, attivando possibilmente la consapevolezza e l'intelligenza emotiva dei singoli.

Tutti questi processi possono essere indirizzati, in generale, allo sviluppo della persona umana ed al controllo di quanto ostacola tale sviluppo, negli individui e nelle comunità. O di quanto può rovesciarlo e annichilirlo, come gli sviluppi distorti dell'IA. Con un orientamento crescente degli Stati, delle comunità, delle organizzazioni, attraverso l'impegno delle

persone che le rappresentano, ciascuno nel proprio ambito tenendo conto della propria situazione, ma collegati dall'impegno comune. Costruendo una grande rete, anche utilizzando strumentalmente sistemi controllati di IA.

Teilhard de Chardin((Nel suo saggio più celebre *"Le Phénomène humain"* Parigi 1955.)) ha definito "noosfera" il livello di coscienza collettiva che si aggiunge alla "biosfera" del mondo vivente grazie alla nascita del pensiero umano e alla connessione spirituale tra gli uomini, ed ha sostenuto che «*Niente nell'universo potrebbe resistere a un numero sufficientemente grande di intelligenze raggruppate e organizzate*». E ben orientate.

In questa prospettiva, la consapevolezza, l'intelligenza razionale, l'intelligenza profonda individuale e collettiva possono produrre un intellettuale collettivo di nuova generazione, a livello globale, in grado di affrontare la policrisi in atto anche sull'aspro versante del controllo dell'IA.

Antonio Zucaro

1. Mons. Vincenzo Paglia, presidente Pontificia Accademia per la Vita, su *AVVENIRE*, settembre 2026.[↔]
2. *"Superintelligence: Paths, Dangers, Strategies"*, saggio del 2014.[↔]
3. *"Il sé digitale. Dai neuroni specchio alla mediazione tecnologica"* Raffaello Cortina Editore, 2026.[↔]
4. Raul Gabriel su *Avvenire*, settembre 2026.[↔]
5. Enciclica *"Magnifica Humanitas"*, 113.[↔]
6. Sia consentita l'autocitazione: Antonio Zucaro *"Buddha, Marx e l'Apocalisse lenta"*, Bordeaux ed., 2026.[↔]
7. *Avvenire*, settembre 2026.[↔]
8. Daniel Goleman *"Intelligenza Emotiva"*, Rizzoli 1996.[↔]