

INTELLIGENZA UMANA VS. INTELLIGENZA ARTIFICIALE: UNA PROSPETTIVA TEOLOGICA

HUMAN INTELLIGENCE VS. ARTIFICIAL INTELLIGENCE: A THEOLOGICAL PERSPECTIVE

Maurizio Buioni C.P.

Dottorato in Teologia (PUG) e Filosofia (UniPG)
mauriziocp@gmail.com
ORCID: 0009-0002-1832-1366

Ricevuto: 17 di settembre, 2024 / Accettato: 28 di ottobre, 2024
DOI: 10.5281/zenodo.14283056

Sommario: L'articolo esplora le sfide che l'intelligenza artificiale (IA) e la robotica pongono alla teologia cristiana, concentrandosi su temi come l'origine, la creazione e i limiti. Da un punto di vista teologico la produzione di macchine IA non può essere equiparata alla creazione dell'uomo, il quale possiede una dignità intrinseca basata sulla sua trascendenza e destinazione alla comunione con Dio. La teologia cristiana stimola a comprendere le dinamiche del rapporto tra uomo e macchina, che potrebbero trasformare radicalmente il nostro modo di vivere, mantenendo sempre un discernimento vigile e oggettivo del bene e del male nelle azioni umane. In questo contesto, la filosofia della persona può affiancare la teologia: essa cerca infatti di definire i confini tra persona e macchina attraverso la ragione, complementare alla fede; propone una visione antropocentrica e non tecnocentrica, e afferma la dignità della persona umana, che non può essere ridotta a una macchina o a semplici funzioni.

Parole chiave: Intelligenza Artificiale (IA). Robotica. Teologia cristiana. Dignità umana. Trascendenza. Rapporto uomo/macchina. Discernimento. Filosofia della persona. Antropocentrismo. Tecnocentrismo.

Abstract: The article explores the challenges that artificial intelligence (AI) and robotics pose to Christian theology, focusing on themes such as origin, creation, and limits. From a theological perspective, the production of AI machines cannot be equated with the creation of man, who possesses an intrinsic dignity based on his transcendence and destiny for communion with God. Christian theology encourages understanding the dynamics of the relationship between man and machine, which could radically transform our way of living, always maintaining a vigilant and objective discernment of good and evil in human actions. In this context, the philosophy of the person can complement theology: it seeks to define the boundaries between person and machine through reason, which is complementary to faith; it proposes an anthropocentric and not technocentric vision, and affirms the dignity of the human person, who cannot be reduced to a machine or mere functions.

Keywords: Artificial intelligence (AI). Robotics. Christian theology. Human dignity. Transcendence. Human/machine relationship. Discernment. Philosophy of the person. Anthropocentrism. Technocentrism.

Le innovazioni scientifiche stanno rapidamente trasformando il futuro dell'umanità, sollevando domande cruciali sul ruolo dell'uomo e di Dio. L'evoluzione tecnica rischia di subordinare l'esistenza umana alla logica dei dispositivi, invertendo il rapporto tra uomo e macchina. La crescita dell'intelligenza artificiale (IA) e delle sue applicazioni pone interrogativi fondamentali: le macchine potranno mai pensare e agire come gli esseri umani? Con l'avanzare della bio-informatica e della robotica, si prospetta un futuro in cui la natura umana potrebbe essere radicalmente alterata, sostituendo corpo e mente con componenti meccanici e informatici. In questo breve articolo, senza alcuna pretesa di essere esaustivi, esploreremo l'intelligenza artificiale, la persona e la natura umana, confrontando prospettive tecnocentriche e teologiche, per trarre infine un bilancio.

Uomo e Macchina: un dialogo sul futuro

La tecnologia, la robotica, il futuro dell'uomo e quello di Dio è una prospettiva di riflessione avvincente:

L'odierna evoluzione della capacità tecnica produce un incantamento pericoloso: invece di consegnare alla vita umana gli strumenti che ne migliorano la cura, si corre il rischio di consegnare la vita alla logica dei dispositivi che ne decidono il valore. Se fino ad oggi la tecnica era al servizio dell'umano, oggi il rischio è che la tecnica prenda il sopravvento e si sostituisca in qualche modo all'umano. Questo rovesciamento è destinato a produrre esiti sconvenienti, la macchina non si limita a guidarsi da sola ma finisce per guidare l'uomo, riducendo così la ragione umana a una razionalità alienata degli effetti che non può essere considerata degna dell'uomo e del resto è già reale il rischio che l'uomo venga tecnologizzato, invece che la tecnica umanizzata. Alle cosiddette macchine intelligenti vengono frettolosamente attribuite capacità che sono propriamente umane¹. (Francesco, 2019b)

La rapida evoluzione delle tecnologie dell'intelligenza artificiale (IA), negli ultimi anni, è caratterizzata da un progresso dirompente, a causa

¹ Nella Lettera *Humana communitas*, del 6 gennaio 2019, papa Francesco rimarca chiaramente il tema delle "tecnologie emergenti e convergenti". La capacità di intervenire sulla materia vivente a livelli sempre più microscopici, di gestire quantità crescenti di informazioni e di monitorare e manipolare i processi cerebrali legati all'attività cognitiva e deliberativa ha conseguenze enormi. Queste tecnologie sfidano i confini della specificità biologica e della differenza spirituale dell'essere umano. In questo contesto, afferma che «la differenza della vita umana è un bene assoluto», (n. 4).

della complessità, ampiezza delle applicazioni e velocità, quantità di dati (*big data*) e allo sviluppo di algoritmi. L'intelligenza artificiale comprende tutte le macchine che simulano determinati aspetti dell'intelligenza umana, quali ragionare e prendere decisioni con crescenti gradi di indipendenza e autonomia. L'obiettivo è ambizioso: progettare macchine che imitano l'uomo fino a sostituirlo. Emerge l'interrogativo: potrà una macchina agire e pensare come l'uomo?

I nuovi sviluppi dalla Intelligenza Artificiale (IA) classica, che si concentra sulla costruzione effettiva nel mondo reale di agenti artificiali in grado di eseguire comportamenti intelligenti, alla Intelligenza Artificiale Incorporata, che aspira a simulare in un mondo virtuale competenze cognitive naturali di alto livello, tipicamente ascritte alla mente umana, sempre più mirano a riprodurre tratti di comportamenti umani o a far interagire l'uomo e la macchina non solo in vista di un potenziamento dell'umano, ma anche per una sua possibile sostituzione, se non superamento. Per non dire dei tentativi percorsi dalla bio-informatica di creare vita virtuale e reale, dove si fa riferimento esplicito a parole-concetti come quello di creazione. Una parte quindi dei sostenitori della IA nel suo insieme, quella cosiddetta forte, si pone l'obiettivo ambizioso di progettare macchine simili all'uomo.

L'introduzione di tecnologie per sostituire parti del corpo e il corpo nella sua integralità, la sostituzione artificiale del pensiero (con le tecnologie informatiche, che consentono di costruire una macchina che riproduca, tramite dei circuiti integrati, le stesse identiche funzioni del cervello umano), la creazione di robot, automi con corpo-macchina e mente-computer (secondo i nuovi scenari delineati dalla robotica), seguendo l'idea lanciata dai teorici dell'intelligenza artificiale o teoria computazionale della mente che riduce il funzionamento del cervello al funzionamento del computer, apre *tecnoprofeticamente* un orizzonte futuro che potrebbe portare a una alterazione radicale della natura umana. Si tenta, infatti, di mettere in correlazione il corpo umano (materia organica) con i computer (materia inorganica), sino alla totale artificializzazione dell'umano sostituendo corpo e mente con sussidi meccanici ed informatici.

È prospettare un futuro della produzione di un uomo-macchina, promettendo immortalità e perfezione illimitata. Una progettazione tecnoscientifica che apre scenari di *oggettificazione* dell'umano, pretendendo il superamento dell'*homo sapiens*. L'impianto nel corpo umano di materiali che con-crescono organicamente, determinando una sorta di fusione

produttiva uomo-macchina, la costruzione di corpi saturi di protesi efficienti, la progettazione di robot autonomizzati delineano una nuova possibile oggettivazione tecnica della natura umana.

La teoria funzionalistica della persona che riduce epistemologicamente la persona all'esercizio delle funzioni cognitive e volitive, ha aperto a tematizzare la disincarnazione del soggetto: la persona è identificabile nell'esistenza di un individuo in grado di esercitare determinate funzioni a prescindere dall'esistenza biologica di un corpo e dell'appartenenza ad una determinata specie. È la previsione di H. Moravec (1999) e di N. Bostrom (2005) che ritengono che la filosofia debba sostenere tali nuovi percorsi, riconoscendo semmai come unico limite a tali sviluppi tecnologici ed informatici i problemi economici. In tale prospettiva è giustificato il diritto a creare *cyborg* in quanto lo si considera un miglioramento della natura dell'uomo; semmai l'unico problema potrà essere quello di evitare la discriminazione tra uomini e trans-uomini.

Il pensiero post-umanista e trans-umanista si richiama alla concezione del dualismo antropologico, che da Platone attraverso Cartesio ha caratterizzato parte della riflessione occidentale. La considerazione del corpo come peso (Platone parlava di tomba dell'anima), la separazione del corpo dalla mente (*res cogitans* e *res extensa* in Cartesio) sono alla base di tale orientamento di pensiero che, a partire da una svalutazione della corporeità, giustifica la fungibilità del corpo oggettificato, ridotto e riducibile consentirà il superamento dei limiti (identificati nel corpo). Il corpo viene ridotto alla condizione di puro oggetto, modificabile e manipolabile, programmabile e riprogrammabile, assoggettato ad un controllo di qualità.

Molti sono profondamente preoccupati per il rapido progresso tecnologico in questo settore. Questi timori si collocano nell'ambito del "biocatastrofismo tecnofobico", che si oppone al "bioprofetismo tecnofilo". Questo punto di vista critica le biotecnologie per il loro potenziale distruttivo, richiamando il pensiero di Nietzsche, svalutando le biotecnologie e celebrando il corpo umano nella sua forma naturale (Lecourt, 1997, 2001 e 2003).

Di fronte a questi scenari futuri possibili una riflessione teologica cristiana e filosofica della persona nella prospettiva ontologica (non funzionalista) può offrire un percorso di riflessione che giustifica i limiti delle biotecnologie informatiche, senza esaltare la tecnologia disprezzando il corpo né esaltare il corpo disprezzando la tecnologia: l'obiettivo è quello di consentire interventi biotecnologici e informatici sul corpo umano senza snaturarne l'identità, evitando trasformazioni irrimediabili della corporeità umana e della

mente umana. La teologia cristiana e la filosofia della persona offrono una prospettiva sull'uomo che ne mostra e giustifica la non equiparazione alla macchina, la sua centralità e dignità².

Origine e limite, persona e natura umana

Le frontiere della ricerca scientifica e tecnologica sollevano domande radicali per la teologia, toccando temi come le origini e il compimento cristologico, che rivelano il disegno divino e la questione del male. La teologia cattolica del Novecento ha rivisto la concezione della creazione e della natura, superando l'idea di una natura autonoma e trovando una nuova convergenza con la scienza e la filosofia. Le tribù ebraiche hanno sviluppato una visione unica del Dio dell'alleanza, riconoscendo una vita intelligente con responsabilità nel creato.

Questa prospettiva avvicina all'origine della vita umana e ai limiti del trans-umano, sollevando la questione dei confini tra umano e divino. Gli studi esegetici mostrano che gli autori biblici hanno evoluto dall'enotheismo al monoteismo, affermando la sovranità di Dio e combattendo l'idolatria. La convergenza tra teologia, scienza e filosofia ha portato a una comprensione integrata dell'origine e dei limiti umani, sottolineando la responsabilità e la trascendenza nella vita intelligente.

La sovranità di Dio, interpretata come trascendenza e immaterialità, si è confrontata con l'evento Cristo, introducendo una struttura trinitaria che unisce divino e umano. Nei secoli, la comprensione di Gesù Cristo ha confermato l'unione tra Dio e uomo, permettendo una creazione libera e autonoma. L'uomo, frutto di un atto sovrano di Dio, è chiamato a curare il creato, rispettando il limite tra umano e divino.

La nozione di partecipazione, rielaborata dal pensiero cristiano dal platonismo, spiega l'autonomia del creato e il rapporto di indipendenza dal Creatore. Dio, come causa suprema, permette a ogni creatura di partecipare dell'essere senza esserne una parte, risolvendo la visione panteista e chiarendo che Dio è l'essere di ogni cosa come causa, non come costitutivo essenziale.

² Vedi Tosolini (2022), Staffolani (2022). Vedi inoltre Kasper (1997), Sanna (2001 e 2004), Benanti (2018), Barone (2020).

In sintesi, la convergenza tra teologia, scienza e filosofia ha portato a una comprensione più integrata dell'origine e dei limiti umani, sottolineando la responsabilità e la trascendenza nella vita intelligente.

La rivelazione collega la signoria umana sulle creature alla signoria divina sul creato. L'umanità apprende la propria autorità attraverso l'esperienza del Dio della liberazione, fondata sulla comunione con Dio, che eleva l'uomo sopra la natura e le gerarchie sociali, giustificando la dignità intrinseca di ogni persona³.

La risurrezione di Cristo rappresenta il compimento del destino umano, proteggendo l'umanità e il creato dalle minacce. La chiave dell'esistenza umana è nelle mani di Dio, non dell'uomo, che non può autodistruggersi. Tuttavia, la storia è segnata da violenza e morte, con il male radicato nell'uomo. Il cristianesimo vede il peccato come una condizione universale, non solo come atti individuali. Sant'Agostino ha individuato la radice del male nell'arbitrarietà umana, dove l'uomo cerca di uguagliarsi a Dio, perdendo il senso del dono e della libertà⁴.

Questa dinamica si manifesta nella condanna e uccisione di Gesù, con Dio che interviene per salvare. L'autonomia dell'uomo include la possibilità di trasgredire, innescando il male. Secondo le Scritture, l'uomo crede erroneamente di sapere cosa porta alla felicità, ignorando che "il salario del peccato è la morte" (Rm 6, 23). La rivelazione cristiana illumina l'iniquità, mostrando che il peccato è una condizione universale. La rivelazione cosmica di Dio insiste sull'ascolto del linguaggio divino nel creato, e la ricerca scientifica e tecnologica riflette la complessità della vicenda umana.

La produzione di macchine di IA non è paragonabile all'atto creativo divino che conferisce dignità all'uomo. Tuttavia, esplorare aspetti sconosciuti delle creature è utile e necessario, nonostante i rischi. La destinazione dell'uomo alla comunione con Dio rende la sua vita inviolabile e giustifica la dignità intrinseca di ogni persona.

³ In questo contesto la dichiarazione *Dignitas infinita* del Dicastero per la Dottrina della Fede, pubblicata il 2 aprile 2024, affronta il concetto di dignità umana in modo approfondito e articolato. Essa sottolinea che ogni essere umano possiede una dignità intrinseca e inalienabile, fondata sulla creazione a immagine e somiglianza di Dio e sulla redenzione in Cristo Gesù. Per il nostro tema rimando ai nn 60-62.

⁴ Sant'Agostino discute la radice del male e l'arbitrarietà umana in diverse opere, ma due delle più rilevanti sono: *Le Confessioni*: dove tra l'altro esamina come il desiderio umano di uguagliarsi a Dio e di esercitare il libero arbitrio senza riconoscere la propria dipendenza da Dio conduca al peccato e al male e *La città di Dio*, in cui, contrapponendo la "Città di Dio" alla "Città dell'uomo", discute come l'orgoglio e l'arroganza umana, che portano l'uomo a cercare di essere come Dio, siano alla radice del male e della corruzione nella società.

Il compito etico principale è discernere ciò che si nasconde nel cuore umano e si manifesta nelle sue opere. La teologia cristiana riflette sul rapporto uomo-macchina, concentrandosi su creazione e arbitrio. La filosofia, complementare alla teologia, cerca di delimitare il confine tra persona e macchina.

La domanda è: una macchina può pensare come un uomo? Tecnicamente possibile, ma la filosofia si chiede se l'uomo sia riducibile a una macchina. La questione morale è: se una macchina può diventare simile all'uomo, è giusto che lo diventi?

La prospettiva tecno-centrica

I tecnofili, sostenitori della tecnologicizzazione del corpo e della mente, immaginano un futuro in cui l'artificiale e il naturale si confondono, eliminando la differenza tra uomo e macchina. Promuovono lo sviluppo di tecnologie convergenti e della robotica/IA per superare i limiti biologici e cognitivi umani, aspirando a una perfezione trans-umana o post-umana.

Questa visione presuppone una concezione riduzionistica e funzionalista dell'uomo, visto come un insieme di parti modificabili e sostituibili tecnologicamente (Benanti, 2012). Il corpo diventa un insieme di componenti meccaniche, mentre la mente è ridotta a elementi neurali digitalizzabili. La robotica e l'IA sostituiscono le funzioni umane con quelle artificiali.

Il funzionalismo sostiene che ciò che conta per identificare una persona sono le funzioni esercitate, indipendentemente dalla natura dell'organismo. La natura umana perde così il suo valore ontologico, e ciò che importa sono le funzioni esibite, sia da un uomo che da una macchina

Materialismo, riduzionismo e funzionalismo caratterizzano una prospettiva non cognitivista, che limita la conoscenza della ragione ai soli fatti empiricamente verificabili o falsificabili. Questa visione riprende il dualismo corpo/mente di origine platonico-cartesiana, dove il corpo biologico è visto come un peso da superare tramite la tecnologia. L'evoluzione tecnologica è considerata una scorciatoia per migliorare l'umano e la specie, diventando un

imperativo la cui omissione è vista come colpa e negligenza (Harris, 2007)⁵.

Si prospetta un futuro tecno-profetico in cui la natura umana potrebbe essere radicalmente alterata, con corpo e mente sostituiti da sussidi meccanici e informatici. La convergenza delle tecnologie porterebbe a una fusione tra umano e artificiale. Tuttavia, questi scenari sono speculativi.

Il transumanesimo, teoria attuale nel dibattito, promuove l'abbandono del biologico a favore del virtuale/artificiale/digitale per espandere le capacità umane e migliorare la vita. Il transumanesimo aspira a un perfezionamento sovra-umano, sia moderato (aumento di bellezza, resistenza fisica e aspettativa di vita) sia radicale (superamento della condizione umana percepita come limite). La Massima Centrale del Transumanesimo (MCT) afferma che è etico e desiderabile usare mezzi tecnoscientifici per superare la condizione umana. In questa visione, il potenziamento umano si realizza nella "condizione tecno-umana" attraverso l'evoluzione tecnologica⁶.

Dal potenziamento delle funzioni umane si passa al potenziamento dell'uomo inteso come uomo potenziato (*enhanced human*) post-umano o uomo come macchina (Marchesini, 2004; Pirni, 2016). È l'orizzonte teorico che tenderà a svuotare i corpi umani in carne ed ossa, riducendoli a componenti biotecnologiche meccaniche ed elettroniche, a flussi di informazioni mutanti in grado di assistere fino a sostituire i processi vitali dell'organismo promettendo una perfezione illimitata. Ciò che conta è che la funzione funzioni il più perfettamente possibile, il corpo non è necessario, se non come involucro della funzione. Nella convergenza delle

⁵ L'autore sostiene che il miglioramento umano attraverso la biotecnologia non solo è moralmente difendibile, ma in alcuni casi è anche moralmente obbligatorio. Tali interventi possono permetterci di vivere vite più lunghe, più sane e più felici, fornendoci ad esempio l'immunità contro malattie come il cancro e l'HIV/AIDS. Sfidando il pensiero convenzionale afferma che l'evoluzione umana dovrebbe essere guidata deliberatamente attraverso la selezione biotecnologica per migliorare le nostre capacità mentali e fisiche.

⁶ Alcuni testi che esplorano le diverse radici filosofiche e culturali del transumanesimo: *Engines of Creation* di K. Eric Drexler: Questo libro del 1986 esplora le applicazioni future della nanotecnologia e della produzione molecolare, influenzando profondamente il pensiero transumanista. *The Phenomenon of Man* di Pierre Teilhard de Chardin è un'opera che combina scienza e spiritualità, proponendo una visione evolutiva dell'umanità che ha ispirato molti transumanisti. *Also sprach Zarathustra* di Friedrich Nietzsche, sebbene non direttamente transumanista, il concetto di *Übermensch* (superuomo) ha influenzato l'idea di trascendere i limiti umani. *The Singularity is Near* di Ray Kurzweil, un libro che prevede un futuro in cui l'intelligenza artificiale supererà quella umana, tema centrale nel transumanesimo. *Transhumanism: Engineering the Human Condition*, di Roberto Manzocco, offre una panoramica storica, filosofica e scientifica del transumanesimo, discutendo le sue radici culturali e le principali questioni etiche. Questi testi forniscono una base solida per comprendere le diverse influenze che hanno contribuito alla formazione del pensiero transumanista.

tecnologie è l'uomo stesso che diventa tecnologia in una sorta di «pan-tecnologismo» (Viola , 2010; Grion, 2012)⁷.

In questo contesto *Transhumanisme: la grande illusion*, di Bernard Vergely critica il movimento transumanista, che sostiene l'uso della tecnologia per migliorare le capacità fisiche e cognitive degli esseri umani. Egli mette in luce diverse problematiche legate a questa visione: *illusione di onnipotenza*: il transumanesimo alimenta l'illusione che l'uomo possa diventare onnipotente e immortale attraverso la tecnologia, ignorando i limiti intrinseci della condizione umana. *Perdita di umanità*: l'adozione indiscriminata delle tecnologie transumaniste potrebbe portare a una perdita di ciò che ci rende umani, come la vulnerabilità, l'empatia e la capacità di vivere esperienze autentiche. *Questioni etiche*: l'autore solleva preoccupazioni etiche riguardo alla disuguaglianza che potrebbe derivare dall'accesso alle tecnologie avanzate, creando una divisione ancora più marcata tra chi può permettersi tali miglioramenti e chi no. *Spiritualità e senso della vita*: il transumanesimo può minare la dimensione spirituale dell'esistenza umana, riducendo la vita a un insieme di processi meccanici e tecnologici. Il testo è dunque un invito a riflettere criticamente sulle promesse del transumanesimo e a considerare le implicazioni profonde che queste tecnologie possono avere sulla nostra identità e sul nostro futuro (Vergely, 2019).

Questi scenari potrebbero essere fantasie sciocche o prognosi verosimili, tardivi bisogni escatologici oppure inediti modelli fantascientifici. Tuttavia, anche se speculativi, richiedono una riflessione critica per comprenderne le problematiche. Non si tratta di aderire a un tecno-catastrofismo tecno-fobico, ma di sviluppare una riflessione filosofica sull'umano, il suo significato e valore, e i limiti delle tecnologie.

L'obiettivo non è esaltare la tecnologia a discapito dell'uomo o viceversa, ma permettere interventi sull'umano senza snaturarne l'identità, evitando trasformazioni irreversibili. La tecnologia non è un destino inevitabile, ma una costruzione umana. Dobbiamo quindi considerare non solo cosa resta umano con le tecnologie, ma soprattutto cosa deve restare umano.

⁷ Grion (2012) esplora il dibattito sul postumanesimo che intende scommettere sulla possibilità di migliorare le capacità fisiche e cognitive dell'essere umano attraverso le tecnologie avanzate come la genetica, le nanotecnologie e la robotica.

La prospettiva umano-centrica

Nonostante i progressi nella creazione di macchine che imitano l'organizzazione cerebrale umana, siamo ancora lontani dal raggiungere un'intelligenza artificiale che riproduca tutte le caratteristiche dell'intelligenza umana. Le macchine attuali possono eseguire compiti complessi, come la raccolta e l'archiviazione di informazioni e la capacità di calcolo, spesso superando le capacità umane in questi ambiti. Tuttavia, la mobilità dei robot rimane una sfida costosa sia finanziariamente che energeticamente.

I tecno-ottimisti credono che sia solo una questione di tempo prima che le macchine possano sostituire ogni funzione umana, superando i limiti fisici e cognitivi. Ma esistono differenze qualitative tra intelligenza artificiale e umana? Le macchine potranno mai sostituire completamente l'uomo? E se fosse possibile, sarebbe auspicabile?

La questione riguarda la possibilità e l'opportunità di creare macchine che possano sostituire l'intelligenza umana in ogni ambito. Esistono ragioni antropologiche, ontologiche e morali per difendere l'unicità e la non sostituibilità dell'intelligenza umana. Alcuni aspetti rendono difficile immaginare che le macchine possano un giorno sostituire completamente l'intelligenza umana.

Una dimensione umana che manca alle macchine è l'autoconsapevolezza. Le macchine non sono consapevoli di sé e non hanno un'identità che permane nel tempo. Inoltre, la filosofia occidentale ha riconosciuto l'importanza di affetti, emozioni e sentimenti nella cognizione umana. Le emozioni influenzano le decisioni umane, un aspetto che l'intelligenza artificiale non può replicare, limitando così il sogno dell'IA forte⁸.

Per imitare integralmente l'intelligenza umana, una macchina dovrebbe includere l'influenza degli stati emotivi sulla cognizione e sulla decisione. Questo richiederebbe un confronto tra informatici, filosofi, psicologi e teorici

⁸ Federico Faggin (2023) esplora diversi temi legati alla coscienza e alla natura umana. *Coscienza e Intelligenza Artificiale*: la coscienza umana è un fenomeno irriducibile che non può essere replicato da macchine o algoritmi, inoltre è intrinsecamente legata alla nostra esperienza soggettiva e non può essere ridotta a semplici processi computazionali. *Limiti dei Computer*: nonostante i progressi nell'intelligenza artificiale, i computer mancano della capacità di comprendere e di avere esperienze soggettive. La loro intelligenza è limitata a processi algoritmici e non può emulare la profondità della coscienza umana. *Natura Umana*: l'autore esplora la natura della vita e della coscienza, suggerendo che c'è un aspetto spirituale e trascendente nell'essere umano che va oltre le spiegazioni scientifiche e tecnologiche. *Esperienze Personali*: il testo include anche le sue riflessioni personali inerenti la sua carriera e le sue esperienze, offrendo una prospettiva unica su come la sua visione della coscienza si è evoluta nel tempo.

della mente. La connessione tra dimensione cognitiva ed emotiva evidenzia la complessità della soggettività umana, che non può essere ridotta a oggettività (Damasio, 1995)⁹.

Un problema emergente dell'IA riguarda la distinzione tra sintassi e semantica. I calcolatori possono operare sui legami sintattici tra simboli, ma non possono interpretare i significati, ponendo un ulteriore ostacolo alla realizzazione di un'IA che imiti completamente l'intelligenza umana.

La semantica non è riducibile alla sintassi, e anche con potenze di calcolo sempre più avanzate, le macchine non possono trattare gli aspetti semantici complessi come fa l'intelligenza umana. Il ragionamento umano si basa su concetti e significati, che derivano dall'esperienza reale. Per creare macchine pensanti, non basta la computazione sintattica; servirebbe anche la capacità di interpretare e fare esperienza come un essere umano, includendo emozioni e sensazioni.

Inoltre, le macchine mancano di motivazione intrinseca. La motivazione umana a cercare la verità, conoscere, inventare e creare è unica. Una macchina pensante dovrebbe non solo risolvere compiti assegnati dall'uomo, ma anche autoassegnarsi compiti e aspirare alla crescita della conoscenza. Attualmente, le reti neurali possono eseguire operazioni intelligenti e auto-modificarsi per migliorare, ma la motivazione resta un input umano.

Ciò che manca alla macchina e resta proprio solo dell'umano è la *possibilità di percepire se stesso come un io e di essere riconosciuto dagli altri come un tu* (D'Agostino, 1999). L'uomo è in grado di percepire la relazionalità come un essere, come una condizione ontologica e antropologica e come un dover essere, come una condizione etica e normativa. In questo senso si giustifica l'eccezionalità dell'uomo rispetto agli altri esseri viventi e alle macchine: solo all'uomo va riconosciuta una dignità intrinseca in quanto è l'unico essere in grado di riconoscere il suo dovere relazionale, dunque è in grado di agire moralmente (rispetto agli altri esseri, simili e dissimili da lui). In questo senso l'uomo, e non la macchina, è propriamente autonomo, in quanto in grado di darsi una norma morale, mentre la macchina può al massimo essere imprevedibile, mediante gli algoritmi.

⁹ In questo lavoro, Damasio sfida la tradizionale separazione cartesiana tra mente e corpo, sostenendo che le emozioni giocano un ruolo cruciale nei processi cognitivi e decisionali e utilizza casi clinici e ricerche neurologiche per dimostrare che le emozioni non sono separate dalla ragione, ma piuttosto essenziali per il funzionamento cognitivo. Un esempio famoso trattato nel libro è il caso di Phineas Gage, che subì una grave lesione cerebrale che alterò drasticamente la sua personalità e capacità decisionale.

Le dimensioni emotiva-emozionale, esperienziale, interpretativa, intuitiva e intenzionale-motivazionale, insieme alla autoconsapevolezza, alla libertà e alla relazionalità interpersonale, sono elementi e segni che mostrano la unicità umana, che evidenziano la differenza uomo-macchina. Una differenza che esiste ora di fatto, e che deve rimanere, di principio, per difendere la unicità dell'umano. Nella misura in cui identifichiamo la specificità umana e le *dimensioni insostituibili* dell'umano, porremo dei *limiti* alla possibilità tecnologica di costruire robot androidi o macchine pensanti: un limite al dominio tecnologico e all'aspirazione di rifare) la natura, in quanto la tecnologia può anche disumanizzare e deumanizzare (Casonato, 2021; Kaplan, 2004). L'umano va preservato, perché conta anche nell'epoca delle macchine.

Anche se gli esseri umani costruiscono AI, selezionano dati, elaborano algoritmi, addestrano macchine, devono mantenere il controllo e la supervisione su ciò che progettano, programmano, applicano; le macchine dovrebbero rimanere un "supporto" alle decisioni umane, che cognitivamente "assistono" le decisioni umane, ma non le sostituiscono.

Le macchine non dovrebbero "competere", ma "completare" le azioni umane. In questo senso anche il linguaggio va mantenuto antropocentrico, considerando le macchine di IA come "automatiche" piuttosto che "autonome" nell'apprendimento. Il termine "autonomia" non può essere applicato ad artefatti, anche se sistemi complessi o cognitivi molto avanzati. La terminologia dei sistemi autonomi è tuttavia ampiamente utilizzata nella letteratura scientifica e nel dibattito pubblico per riferirsi al più alto grado di automazione e al più alto grado di indipendenza dagli esseri umani in termini di operazioni e decisioni (casonato & Marchetti, 2021).

La necessità di mantenere il controllo umano resta essenziale anche per evitare il possibile problema della delega tecnologica. Un sistema esperto che diventa ottimale nel suggerire "decisioni" all'uomo pone il rischio di diminuire l'attenzione umana con la possibile conseguenza di ridurre le abilità umane (il cosiddetto fenomeno di de-professionalizzazione), riducendo la responsabilità andando verso l'artificializzazione delle scelte che possono impoverire e perfino cancellare il rapporto interpersonale. In questo senso, è importante disciplinare la sinergia produttiva come complementarità tra uomo e macchina, ricercando modalità di sostegno intelligente che consentano all'uomo di avere «un controllo umano significativo o significativo» (*meaningful human control*) in termini di attenzione, contributo, controllo. e responsabilità. Il dibattito

sulla affidabilità, spiegabilità, tracciabilità, trasparenza e algoritmica ne sono un importante esempio (Floridi et al., 2018).

Osservazioni conclusive

Rifondare un umanesimo tecnologico significa evitare da un lato di *antropomorfizzare* la macchina (umanizzare la tecnologia) o *artificializzare* l'uomo a macchina biologica (tecnologizzare l'uomo). La *macchina* può essere una alleata dell'uomo e l'integrazione tra naturale ed artificiale può divenire complementare, sempre nella priorità del rispetto dell'umano, salvaguardando ciò che conta dell'umano.

La prospettiva cristiana ci insegna a prendere le distanze da uno sviluppo incontrollato del progresso scientifico, a causa della finitezza umana. Il peccato originale è un tentativo di diventare come Dio. Quello che è male, secondo la prospettiva cristiana è l'arroganza umana che riguarda il tema del potenziamento quando diventa un modo di giocare a fare Dio, di diventare Creatore (non co-creatore) e sostituire Dio (negando la Sua esistenza). La prospettiva cristiana ci ricorda che non siamo Dio, né infiniti come Dio. Non possiamo sfuggire alla finitezza della creazione. Il tentativo finisce in una tragedia.

In questo senso, il transumanesimo attraverso il potenziamento funzionale umano cerca una sorta di "salvezza tecnologica" nella realizzazione materialistica dell'eterna immortalità sulla terra senza Dio, come nel progetto della Torre di Babele, una forma di *arroganza*, una ribellione contro la condizione umana, una ribellione contro il Creatore, allo scopo di dimostrare di essere migliore di Dio, tecnologicamente. È una prospettiva post-umanista, anti-cristiana o post-cristiana. Nella prospettiva cristiana lo scopo è la salvezza teologica spirituale, nell'accettazione della finitezza e della vulnerabilità come condizioni umane e come dono e nell'identificazione della felicità con la dimensione spirituale "al di là" del miglioramento materialistico e funzionalistico.

All'interno di questa prospettiva si delinea una filosofia della persona umana che, in contrapposizione alla visione funzionalista libertaria ed utilitarista (che riduce la persona all'esercizio di funzioni razionali, volitive e sensitive), ne giustifica la consistenza ontologica: la persona si identifica con

l'essere umano avente natura umana a prescindere dalle funzioni esercitate e dalla loro potenza. Tale orizzonte antropologico evidenzia che la natura umana è ciò che conta nell'era della tecnoscienza quale riferimento che può guidare l'etica nella definizione del rapporto uomo/macchina, nel riconoscimento delle potenzialità e dei limiti, nel rispetto dell'umano e di ciò che va conservato dell'umano per ciò che è e non per ciò che fa.

L'IA è progettata e prodotta da un essere umano: siamo noi che dobbiamo decidere come produrre l'IA. Non sappiamo cosa sarà fattualmente e tecnicamente possibile: ma sappiamo oggi cosa è umano e cosa vogliamo conservare di ciò che è umano in una prospettiva incentrata sull'uomo, ispirata da una visione teologica cristiana e ad una visione della filosofia della persona umana che ne possono segnare l'orizzonte valoriale di riferimento.

La fede cristiana, dunque, offre elementi per la salvezza dell'uomo decaduto, risanandolo e portandolo alla pienezza grazie a Gesù Cristo, Dio fatto uomo. La riflessione teologica ha sempre difeso il ruolo della carnalità nel destino ultimo dell'uomo, fin dai primi secoli della nostra era, quando si è confrontata con lo gnosticismo, che sosteneva una salvezza che veniva dalla conoscenza e che era una liberazione dalla materialità¹⁰.

A chi propone di salvare solo la mente, dobbiamo ricordare che l'elemento fondamentale del cristianesimo è che il *Logos* di Dio si è fatto carne in Gesù Cristo. Inoltre, ricordiamo, tenendo presente il simbolo apostolico, che nel cristianesimo la salvezza finale è associata alla risurrezione della carne, che è molto più di una rivivificazione del corpo.

Inoltre, in contrasto con la pretesa umana di raggiungere la salvezza con le proprie forze, il cristianesimo ha riconosciuto il carattere gratuito di questa salvezza, con figure come Paolo e Agostino¹¹.

Le vecchie tentazioni sono riapparse nel corso della storia e le ritroviamo in alcuni dei fautori dell'IA forte (come H. Moravec e R. Kurzweil). Il fisico viene rifiutato per abbracciare il virtuale. Una salvezza ottenuta con mezzi tecnici, che si suppone ci liberi dalla carne, per ridurre

¹⁰ A questo proposito, spiccano Ireneo di Lione e Tertulliano (160 ca. - 220 ca.). La famosa espressione di Tertulliano, "*caro salutis est cardo*" ("la carne è la pietra angolare della salvezza"), nel *De resurrectione carnis* - o *De resurrectione mortuorum*, VIII, 2 (PL 2, 806; CCL 2, p. 931).

¹¹ Qui troviamo la dialettica paolina tra fede e opere, così come la successiva controversia tra Agostino e Pelagio, nel primo terzo del V secolo, che ha segnato la storia della teologia cristiana occidentale. Lo stesso Papa Francesco, in numerose occasioni, ha denunciato quella che considera la presenza dello gnosticismo e del pelagianesimo nel mondo di oggi (Francesco, 2013 e 2018; Congregazione per la Dottrina della Fede, 2018).

la nostra storia a dati, non sarà una vera salvezza, perché non risponde alla totalità dell'uomo, né al desiderio di pienezza presente nel suo cuore. Anzi, può apparire piuttosto come una condanna. Vediamo qui riattualizzato il film costante dell'antica gnosi.

Di fronte alla sfida tecnologica dell'umanità, la teologia, come antica sapienza che cerca di rivelare la statura della vita umana, chiamata Vita, con l'iniziale maiuscola, ha il compito di ricordare all'uomo la sua origine e il suo destino, senza cedere a illusioni che possono essere fuorvianti. Ancora una volta le parole di Francesco (2024) risuonano forte infatti:

La stagione di innovazione tecnologica che stiamo attraversando si accompagna a una particolare e inedita congiuntura sociale. Si registra come uno smarrimento o quantomeno un'eclissi del senso dell'umano e un'apparente insignificanza del concetto di dignità umana. Ed è così che in questa stagione in cui i programmi di intelligenza artificiale interrogano l'essere umano e il suo agire, proprio la debolezza dell'*ethos* connesso alla percezione del valore e della dignità della persona umana rischia di essere il più grande *vulnus* nell'implementazione e nello sviluppo di questi sistemi. Non dobbiamo dimenticare infatti che nessuna innovazione è neutrale: nessuna innovazione è neutrale. La tecnologia nasce per uno scopo e, nel suo impatto con la società umana, rappresenta sempre una forma di ordine nelle relazioni sociali e una disposizione di potere, che abilita qualcuno a compiere azioni e impedisce ad altri di compierne altre. Questa costitutiva dimensione di potere della tecnologia include sempre, in una maniera più o meno esplicita, la visione del mondo di chi l'ha realizzata e sviluppata.

Riferimenti bibliografici

- Barone, C. (a cura di). (2020). *L'algorithmo pensante. Dalla libertà dell'uomo all'autonomia delle Intelligenze Artificiali*. Il Pozzo di Giacobbe.
- Benanti, P. (2012). *The cyborg: Corpo e corporeità nell'epoca del post-umano*. Cittadella.
- Benanti, P. (2018). *Le macchine sapienti. Intelligenze artificiali e decisioni umane*. Marietti 1820.
- Bostrom, N. (2005). In defense of posthuman dignity. *Bioethics*, 19(3), 202–214.
- Casonato, C. (2019). Potenzialità e sfide dell'intelligenza artificiale. *BioLaw Journal - Rivista di BioDiritto*, 1, Articolo 1. <https://doi.org/10.15168/2284-4503-359>
- Casonato, C., & Marchetti, B. (2021). Prime osservazioni sulla proposta di regolamento dell'Unione Europea in materia di intelligenza artificiale. *BioLaw Journal - Rivista di BioDiritto*, 3, 415-437. <https://doi.org/10.15168/2284-4503-1793>

- Congregazione per la Dottrina della Fede. (2018, febbraio 22). *Lettera Placuit Deo ai Vescovi della Chiesa cattolica su alcuni aspetti della salvezza cristiana (22 febbraio 2018)*. https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_con_cfaith_doc_20180222_placuit-deo_it.html
- D'Agostino, F. (1999). La dignità umana, tema bioetico. In P. Cattorini, E. D'Orazio, & V. Pocar (a cura di), *Bioetica e cinema: Racconti di malattia e dilemmi morali* (Vol. 43, p. 103 e ss.). Zadig.
- Damasio, A. (1995). *L'errore di Cartesio. Emozione, ragione e cervello umano*. Adelphi.
- Dicastero per la Dottrina da Fede. (2024, aprile 2). *Dichiarazione Dignitas infinita circa la dignità umana (2 aprile 2024)*. https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_ddf_doc_20240402_dignitas-infinita_it.html
- Faggin, F. (2023). *Irriducibile. La coscienza, la vita, i computer e la nostra natura*. Mondadori.
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Francesco. (2013, novembre 24). «*Evangelii gaudium*»: *Esortazione Apostolica sull'annuncio del Vangelo nel mondo attuale*. https://www.vatican.va/content/francesco/it/apost_exhortations/documents/papa-francesco_esortazione-ap_20131124_evangelii-gaudium.html
- Francesco. (2018, marzo 19). «*Gaudete et exsultate*»: *Esortazione Apostolica sulla chiamata alla santità nel mondo contemporaneo*. https://www.vatican.va/content/francesco/it/apost_exhortations/documents/papa-francesco_esortazione-ap_20180319_gaudete-et-exsultate.html
- Francesco. (2019a, gennaio 6). *Lettera del Santo Padre Francesco al Presidente della Pontificia Accademia per la Vita in occasione del XXV anniversario della sua istituzione*. https://www.vatican.va/content/francesco/it/letters/2019/documents/papa-francesco_20190106_lettera-accademia-vita.html
- Francesco. (2019b, febbraio 25). *Ai Partecipanti all'Assemblea Plenaria della Pontificia Accademia per la Vita*. https://www.vatican.va/content/francesco/it/speeches/2019/february/documents/papa-francesco_20190225_plenaria-accademia-vita.html

- Francesco. (2024, giugno 14). *Partecipazione del Santo Padre Francesco al G7 a Borgo Egnazia*. <https://www.vatican.va/content/francesco/it/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>
- Grión, L. (a cura di). (2012). *La sfida postumanista. Colloqui sul significato della tecnica*. Il Mulino.
- Harris, J. (2007). *Enhancing Evolution: The Ethical Case for Making Better People*. Princeton University Press. <https://www.jstor.org/stable/j.ctt7t8vz>
- Kaplan, D. M. (2004). *Readings in the Philosophy of Technology*. Rowman & Littlefield.
- Kasper, W. (1997). La Chiesa di fronte alle sfide del postmoderno. *Humanitas*, 52, 171–189.
- Lecourt, D. (1997). *Il mito del progresso*. Editions Textuel.
- Lecourt, D. (2001). *La bioetica è in malafede?* Luca Sossella Editore.
- Lecourt, D. (2003). *Humain post-humain: La Technique et la vie*. PUF.
- Marchesini, R. (2004). *Post-human: Verso nuovi modelli di esistenza*. Bollati Boringhieri.
- Moravec, H. P. (1999). *Robot: Mere Machine to Transcendent Mind*. Oxford University Press.
- Pirni, A. (a cura di). (2016). *Il postumano realizzato. Orizzonti di possibilità e sfide per il nostro tempo*. Interlinea.
- Sanna, I. (2001). *L'antropologia cristiana tra modernità e postmodernità*. Queriniana.
- Sanna, I. (2004). L'antropologia della postmodernità ed il messaggio cristiano: Sfide e prospettive. In R. Nardin (a cura di), *Vivere in Cristo. Per una formazione permanente alla vita monastica* (pp. 59–105). Città Nuova.
- Staffolani, M. (2022). Il Vangelo della Croce e la condizione tecno umana. Tra grandi possibilità e l'inevitabile limite della morte. In *La Sapienza della Croce in un mondo plurale. Atti del IV Congresso Teologico Internazionale. Volume III* (pp. 37–49). Editrice Velar.
- Tosolini, T. (2022). L'atmosfera postmoderna e la grazia della croce. In F. Taccone & C. Benedettini (a cura di), *La Sapienza della Croce in un mondo plurale. Atti del IV Congresso Teologico Internazionale. Volume III* (pp. 9–36). Editrice Velar.
- Vergely, B. (2019). *Transhumanisme: La grande illusion*. Le Passeur.
- Viola, F. (2010). Umano e post-umano: La questione dell'identità. In F. Russo (a cura di), *Natura, cultura, libertà*. Armando Editore.