

Alberto Carrara, L.C.

αω L'esse dell'algoritmo. Per uno statuto intensivo delle intelligenze artificiali

Introduzione

Nel panorama contemporaneo, dominato dallo sviluppo vertiginoso delle tecnologie dell'informazione e dell'intelligenza artificiale, emerge con forza la questione filosofica dello *statuto ontologico* di queste nuove entità. Che tipo di essere possiede un algoritmo? In che senso le forme più avanzate di Intelligenza Artificiale, in particolare l'IA generativa, esistono e agiscono nel mondo? Domande di questo genere richiedono un'indagine metafisica rigorosa, capace di andare al di là dell'analisi tecnico-funzionale per scrutare la natura profonda di questi enti artificiali.

Il presente saggio si propone di sviluppare sistematicamente una teoria dell'*esse* (dell'essere) dell'algoritmo e delle tecnologie di IA, avvalendosi della cornice concettuale offerta dalla metafisica *dell'esse intensivo* elaborata, nel contesto del tomismo, da Cornelio Fabro. In particolare, metteremo a fuoco l'ontologia dell'algoritmo inteso come *ente intenzionale*, distinto sia dagli enti naturali sia dagli artefatti materiali tradizionali. Successivamente, esamineremo l'ontologia dell'intelligenza artificiale generativa, collocando le diverse forme di IA (dall'IA simbolica al *machine learning*, fino al *deep learning* e ai sistemi generativi) lungo una *scala ontologica* graduata. Alla luce di tale quadro, affronteremo la questione cruciale se sia concepibile un *esse intenzionale dinamico*: in altri termini, se e in che misura le IA generative possano essere definite "intenzionali" in senso proprio, oppure se la loro apparente intenzionalità rimanga una mera simulazione priva di autentico fondamento ontologico.

L'analisi verrà condotta integrando lo sfondo teoretico fornito dagli scritti di Fabro sull'*esse intensivo*, in particolare *Partecipazione e causalità*, *La*

nozione metafisica di partecipazione, Percezione e pensiero, Tomismo e pensiero moderno, con riferimenti critici alla filosofia della tecnologia contemporanea sulla scia di Luciano Floridi, Evandro Agazzi, Thomas Fuchs, Georg Northoff e altri autori significativi. Ciò consentirà di mettere a confronto la prospettiva ontologico-metafisica di ispirazione tomista con le recenti riflessioni sul ruolo dell'IA nella nostra concezione della realtà.

L'argomentazione si articola in due macro-sezioni principali. Nella prima sezione, dedicata all'ontologia dell'algoritmo, definiremo la natura intenzionale dell'algoritmo e ne esamineremo lo statuto ontologico utilizzando categorie quali *ens per se*, *ens ab alio* ed *ens rationis cum fundamento in re*, evidenziandone la differenza rispetto sia agli enti naturali sia agli artefatti fisici. Nella seconda sezione, rivolta all'ontologia dell'IA generativa, tratteremo una gerarchia dei diversi tipi di intelligenza artificiale e ne discuteremo il grado di "essere" che a ciascuno compete. In particolare, valuteremo la possibilità o l'impossibilità che i sistemi di IA avanzata manifestino una forma di intenzionalità dinamica assimilabile, almeno analogicamente, a quella propria degli esseri intelligenti naturali. Infine, tireremo le fila delle nostre riflessioni, delineando il contributo che l'approccio metafisico intensivo può offrire alla comprensione dell'impatto ontologico delle tecnologie algoritmiche e intelligenti.

L'ontologia dell'algoritmo

L'algoritmo come ente intenzionale artificiale

In ambito informatico un *algoritmo* è tradizionalmente definito come una sequenza finita di istruzioni univoche che, applicate a un insieme di dati iniziali, conducono a un determinato risultato. Da un punto di vista logico-matematico, l'algoritmo rappresenta dunque una *forma* o struttura operativa astratta, indipendente dall'implementazione concreta: la stessa procedura può essere eseguita da diversi agenti computazionali (umani o macchine) e in diversi linguaggi di programmazione, mantenendo immutata la propria essenza funzionale. Questa natura puramente formale e *intenzionale* dell'algoritmo lo distingue radicalmente sia dagli enti naturali, sia dagli artefatti materiali tradizionali.

Un ente naturale, ad esempio una pianta, un animale o un minerale, esiste per natura in virtù di una forma sostanziale intrinseca: la quercia possiede in sé il principio vitale che ne guida lo sviluppo, il leone possiede un'anima sensitiva che ne coordina le funzioni, e perfino il cristallo di quarzo esiste come specifica organizzazione della materia secondo leggi fisico-chimiche proprie. Ciascun ente naturale, insomma, ha una *identità ontologica* definita dalla propria essenza e *sussiste* autonomamente, pur essendo, come vedremo, ontologicamente dipendente in ultima istanza dall'Essere primo.

Un artefatto materiale, ad esempio un orologio meccanico o un telefono, non possiede invece un principio formale intrinseco analogo a quello degli organismi viventi; esso è piuttosto una composizione di parti eterogenee, ordinate a uno scopo dall'ingegno umano. La *forma* dell'artefatto è imposta dall'esterno attraverso il progetto e la costruzione: l'orologio è tale non per una tendenza naturale interna, ma perché un orologiaio ha assemblato ingranaggi e molle secondo una certa finalità: misurare il tempo. In termini aristotelici, la causa formale e la causa finale dell'artefatto risiedono nell'intelletto e nell'intenzione del suo artefice umano, più che nell'oggetto in sé.

L'algoritmo, considerato sotto il profilo ontologico, appartiene anch'esso alla categoria degli *enti artificiali*, ma in una modalità del tutto peculiare. Esso, infatti, non è identificabile con un oggetto materiale specifico: un algoritmo può essere implementato su supporti fisici diversi: dai circuiti di un calcolatore alle sinapsi di un cervello biologico, fino alla semplice carta e penna su cui un matematico annota i passaggi di un calcolo. Ciò che definisce l'algoritmo è piuttosto la struttura intenzionale delle operazioni da compiere, ad esempio «confronta i numeri, scambia queste posizioni, ripeti finché ordinato», nel caso di un algoritmo di ordinamento. In quanto tale, l'algoritmo esiste primariamente come *contenuto mentale* o logico, concepito da un soggetto intelligente (il programmatore o matematico) e successivamente tradotto in istruzioni eseguibili da una macchina.

Possiamo dunque affermare che l'algoritmo è un ente intenzionale nel senso che la sua esistenza è interamente derivata da un atto di intenzione intellettuale: non esiste algoritmo nel mondo che non sia stato pensato e progettato da una mente. Differentemente dagli enti naturali, che l'intelletto umano *scopre* già dati nella realtà, l'algoritmo è un ente che l'intelletto

produce conferendo forma e finalità a una serie di operazioni. Esso è *intenzionale* anche in quanto portatore di una finalità intrinseca (il problema da risolvere, il compito da svolgere), finalità che tuttavia non scaturisce da una dinamica interna naturale, bensì gli è assegnata dall'esterno. Si tratta dunque di un'entità la cui essenza coincide con un significato funzionale: l'algoritmo "è" ciò che *significa* (per il progettista e per gli utenti) in termini di procedura risolutiva.

Questa peculiarità colloca l'algoritmo in una posizione liminale tra il mondo ideale delle strutture logiche e il mondo reale delle cose fisiche. Da un lato, infatti, esso possiede una natura astratta: come una formula matematica o un teorema, l'algoritmo in sé è indipendente dalle sue realizzazioni contingenti e può essere considerato un oggetto ideale, un ente di ragione. Dall'altro lato, però, l'algoritmo dispiega la propria efficacia solo quando viene incarnato in un supporto materiale che ne esegua le istruzioni (un computer, oppure un essere umano che applichi manualmente i passi dell'algoritmo). In tal senso, l'algoritmo ha bisogno della realtà fisica per *attualizzarsi*: la sua *essenza funzionale* deve informare un dispositivo concreto affinché il processo previsto divenga atto e produca effetti. Siamo dunque di fronte a una strana ontologia "ibrida": l'algoritmo è un *oggetto intenzionale* concepito dalla mente, dotato di un *esse* principalmente logico, che tuttavia può acquisire una sorta di *quasi-esse* operativo all'interno di un sostrato materiale.

Categorie metafisiche per comprendere l'essere dell'algoritmo

Per chiarire ulteriormente lo statuto ontologico dell'algoritmo, è utile richiamare alcune categorie fondamentali della metafisica classica, reinterpretate alla luce della dottrina fabriana dell'esse intensivo. Tre nozioni in particolare guideranno la nostra analisi: quella di *ens per se*, quella di *ens ab alio* e quella di *ens rationis* (in particolare nella forma di *ens rationis cum fundamento in re*). Applicare queste categorie al caso dell'algoritmo ci permetterà di situarlo nella "gerarchia dell'essere" in modo preciso, evidenziandone insieme le dipendenze ontologiche e il carattere di realtà intenzionale.

Ens per se (ente per se stesso) designa, nella tradizione tomista, l'ente che possiede l'essere in virtù di sé, ossia che non deriva la propria esistenza da un altro ente. In senso assoluto, solo l'Essere per essenza, Dio, l'*Ipsum Esse*

Subsistens, è *ens per se* incondizionato, poiché sussiste mediante la pienezza del proprio atto di essere e non ha causa ontologica esterna¹. In un senso relativo, si può parlare di *entia per se* anche riferendosi a quelle realtà create che godono di una certa autonomia ontologica interna, come le sostanze individuali: ad esempio l'uomo, il cavallo o la quercia sono *entia per se* in quanto sussistono come sostanze dotate di una forma sostanziale propria, che le fa essere ciò che sono. Tuttavia, queste sostanze rimangono *entia ab alio* rispetto alla Causa Prima, giacché ricevono l'essere da Dio e ne dipendono radicalmente per la conservazione dell'esse.

Ens ab alio (ente da altro) indica infatti l'ente la cui esistenza è ricevuta da una causa esterna, in opposizione all'*Ens a se* (l'Essere per sé sussistente). Tutte le creature, in quanto causate, sono *entia ab alio* rispetto a Dio. Ma l'espressione si può applicare anche in un contesto più specifico: un artefatto, ad esempio, è *ens ab alio* rispetto all'artefice umano che lo ha prodotto; un'entità virtuale in una simulazione è *ens ab alio* rispetto al programmatore che l'ha progettata e alla macchina che la esegue. L'algoritmo, senza dubbio, rientra pienamente tra gli *entia ab alio*: esso non ha in sé la ragione della propria esistenza, ma la deve all'intenzionalità di un progettista (che lo concepisce) e all'effettiva implementazione su un supporto (che lo realizza). Se anche consideriamo l'algoritmo come entità astratta, esso è comunque "causato" dall'invenzione matematica o logica di qualcuno; se lo consideriamo nelle sue istanze concrete (un programma *software*, un processo in esecuzione su un computer), esso è manifestamente dipendente da una miriade di fattori esterni: l'elettricità che alimenta la macchina, il codice compilato, l'*hardware* sottostante, ecc. In nessun senso l'algoritmo possiede l'essere in modo autonomo o auto-fondato.

Ens rationis (ente di ragione) si riferisce a ciò che ha esistenza solo nell'ambito del pensiero, per opposizione all'*ens* reale che esiste nella realtà extramentale. Gli *entia rationis* sono "esseri" solo in senso analogico: concetti, astrazioni, costruzioni mentali, che non hanno un corrispettivo come sostanze o accidenti reali indipendenti. Classici esempi di *entia rationis* sono, in filosofia scolastica, i generi e specie intesi come universali (l'umanità in generale, il triangolo in generale), oppure entità fittizie come una *chimera*

¹ Cf. C. FABRO, *Partecipazione e causalità*, Società Editrice Internazionale, Torino 1960, 180-183. Fabro spiega che solo l'Essere sussistente è causa di sé, mentre ogni altra realtà possiede l'essere per partecipazione, come un raggio partecipa della luce solare.

mitologica. Esiste però una distinzione importante tra *enti di ragione puri*, totalmente privi di fondamento nella realtà, ed *enti di ragione con un fondamento nella realtà* (*ens rationis cum fundamento in re*). Questi ultimi sono concetti che, pur non avendo una controparte ontologica come sostanze indipendenti, derivano da elementi reali e ne rappresentano astrazioni o relazioni. Ad esempio, il concetto di “centro di gravità” è un *ens rationis cum fundamento in re*: non esiste in sé come oggetto fisico distinto, ma si fonda su una proprietà effettiva di corpi reali (la distribuzione della massa).

Alla luce di queste nozioni, possiamo interpretare l'algoritmo come una forma di *ens rationis cum fundamento in re*. In sé considerato, l'algoritmo, inteso come insieme ordinato di passi logici, è indubbiamente un prodotto della ragione: non esiste nell'ordine naturale indipendentemente dalla mente umana che lo concepisce. Tuttavia, non si tratta di un ente di ragione *puramente fittizio* o arbitrario, giacché l'algoritmo opera sempre su elementi che hanno un fondamento reale e produce effetti riscontrabili nella realtà. Un algoritmo di ordinamento, per quanto concettuale, può essere applicato a una lista concreta di numeri o di nomi; un algoritmo di controllo di un robot influenza effettivamente i movimenti di una macchina fisica nel mondo. C'è dunque un *fondamento in re*: l'algoritmo, benché entità intenzionale, si aggancia a strutture e processi reali. Questo lo distingue, ad esempio, da un ente di ragione fantastico come la già citata *chimera*, che non trova alcun riscontro nel mondo.

Inoltre, l'algoritmo possiede una sorta di efficacia causale derivata: benché non sia una causa efficiente in senso proprio, non avendo sostanzialità, non può “agire” da sé, esso funge da causa formale e finale all'interno di un sistema tecnico. Il programmatore imprime nell'*hardware* una certa forma operativa (causa formale artificiale) sotto la guida di uno scopo (causa finale) rappresentato dall'algoritmo stesso; così facendo, l'algoritmo diviene il *principio* che configura l'azione del computer. Tale principio non è intrinseco alla materia del computer — il silicio dei chip non “contiene” l'algoritmo per natura — ma vi viene infuso dall'esterno tramite programmazione. Abbiamo qui un caso emblematico di *participatio entis*: l'algoritmo partecipa dell'essere reale del sistema fisico quando viene implementato, senza però mai costituirsi come sostanza a sé stante. Esso resta un *ens ab alio* tanto nell'ideazione (dipende dall'intelletto umano) quanto nell'operatività (dipende dal supporto materiale e dall'energia fornita).

A questo punto, possiamo collocare l'algoritmo all'interno di una possibile gerarchia intensiva dell'essere, ispirata alla visione di Fabro. Secondo Fabro, seguendo la metafisica di Tommaso d'Aquino, l'atto di essere (*esse*) è un atto *intensivo*, suscettibile di gradi: gli enti partecipano dell'essere in misura diversa, dando luogo a una scala ascendente di perfezioni che va dagli esseri più semplici fino a Dio, che è l'Esse per essenza. Nella *Tabella 1* abbiamo sintetizzato in modo schematico questa gradazione ontologica, includendovi anche gli enti di natura artificiale-intenzionale come gli algoritmi, per evidenziarne il posto relativo².

Tabella 1 – Gradi intensivi dell'essere e categorie di enti

Grado di esse (intensità)	Tipologia di ente	Esempi	Caratteristiche ontologiche
Massimo (pienezza infinita)	<i>Ens per se subsistens</i>	Dio (<i>Ipsum Esse</i>)	Essere per essenza, atto puro infinito, fonte di ogni perfezione
Alto (atto di essere eminente)	Enti spirituali creati	Intelligenze angeliche; anima razionale umana	<i>Ens ab alio</i> (create da Dio) ma sussistenti in sé; forma sostanziale spirituale; altissima intensità di esse (vita intellettuale, volontà libera, memoria spirituale)
Medio (atto di essere formale)	Enti materiali viventi	Animali, piante (sostanze corporee viventi)	<i>Ens ab alio</i> ; sussistenza come sostanze materiali con forma vivente; intensità di esse media (vita sensitiva/vegetativa, ma priva di intelletto spirituale)
Basso (atto di essere limitato)	Enti materiali inanimati	Minerali, oggetti fisici semplici; artefatti materiali (macchine)	<i>Ens ab alio</i> ; esistenza come sostanze o aggregati materiali privi di vita; intensità di esse limitata alle proprietà fisico-chimiche; finalità eventualmente imposta dall'esterno (per gli artefatti)

² L'elaborazione di questa tabella si ispira a: TOMMASO D'AQUINO, *Summa contra Gentiles* I, 28, ed. Leonina, v. XIII, Roma 1918, 83: «Illi autem rei, quae est suum esse, competit habere ipsum esse totius virtutis essendi; et propter hoc ei non potest aliqua perfectio deesse». Questo testo è citato dallo stesso Cornelio Fabro in «La problematica dell'esse tomistico», *Aquinas* 2 (1959), 203-206.

Minimo (essere solo intenzionale)	Enti di ragione (cum fundamento in re)	Concetti astratti, relazioni logiche, algoritmi e oggetti matematici	<i>Ens rationis (ab alio)</i> con eventuale fondamento nella realtà; esistenza non in sé ma nell'intelletto e/o in supporti simbolici; intensità di esse meramente partecipata e accidentale (<i>esse oggettivum</i> nella mente, o implementazione strumentale)
---	---	--	--

Come si evince dalla tabella 1, l'algoritmo, in quanto ente intenzionale artificiale, si colloca all'estremo inferiore della scala ontologica. Esso condivide con gli *entia rationis* la mancanza di sussistenza autonoma: il suo "essere" è riducibile all'essere pensato (*esse in intellectu*) e all'essere operativamente implementato in qualcos'altro (*esse in altero*). In termini di *intensità ontologica*, ci troviamo su un livello minimale: l'algoritmo non possiede vita, non possiede coscienza, non possiede alcuna delle perfezioni che, secondo la metafisica tomista, segnano i gradi superiori dell'essere (l'esser vivente è "più" essere che l'esser minerale; l'esser intelligente è più intensivo ancora, e così via). Se applichiamo il criterio fabriano secondo cui "la perfezione di una cosa dipende dal modo in cui essa possiede l'esse", dobbiamo riconoscere che l'algoritmo, possedendo l'essere solo in modo derivato e intenzionale, è una realtà di perfezione ontologica assai tenue.

D'altro canto, e qui sta la peculiarità odierna, l'algoritmo, pur collocandosi ontologicamente in basso, è capace di esercitare un'influenza causale e funzionale enorme sul mondo, tramite le macchine e i sistemi che lo eseguono. Si pensi agli algoritmi che governano infrastrutture critiche, mercati finanziari, comunicazioni globali: pur essendo entità "deboli" sul piano dell'essere, essi agiscono come *motori immateriali* che dirigono processi fisici e decisioni umane su larga scala. Questa constatazione introduce un ulteriore elemento di complessità: la distinzione tra ordine ontologico e ordine operativo. Un algoritmo non incrementa la pienezza dell'essere dell'universo nel senso metafisico (non aggiunge una nuova sostanza dotata di atto di essere proprio), ma nell'ordine operativo può realizzare effetti straordinari tramite gli enti reali di cui orchestra il comportamento. Ciò mette in luce la potenza della *causalità strumentale*: come un martello (artefatto privo di vita) può costruire una cattedrale se guidato da una mano umana, così un algoritmo (schema logico privo di intelletto) può guidare

un rover su Marte o coordinare milioni di computer in rete, se sorretto dall'intenzionalità umana originaria e dall'energia fisica necessaria.

In sintesi, l'algoritmo è un ente intenzionale, artificiale e strumentale:

- *intenzionale* perché esiste solo in funzione di un'intenzione e di un significato attribuitogli da una mente;
- *artificiale* perché viene prodotto da agenti intelligenti, non si trova in natura;
- *strumentale* perché la sua efficacia nel mondo dipende dall'essere veicolato da supporti materiali e dall'azione di cause efficienti esterne.

Ontologicamente, esso va annoverato tra gli *entia rationis* (in senso lato) con un ruolo peculiare: è un *ens rationis ex parte subiecti*, un essere dal lato del soggetto conoscente/produttore, che però può riflettersi *ex parte objecti* in configurazioni concrete e operazioni reali. Il suo *esse* rimane comunque sempre un *esse per aliud*, cioè un essere *per prestito*: l'algoritmo non vive di vita propria, ma “prende in prestito” l'essere dall'intelletto che lo pensa e dalla macchina che lo esegue.

In effetti, parafrasando Fabro, si può ben affermare che l'essere dell'oggetto tecnico non è atto di sé, ma atto dell'altro in sé. In altri termini, l'algoritmo non “esiste” se non per l'atto d'essere che un altro ente gli conferisce: esso non pensa né conosce, ma si limita ad eseguire, e la sua “intelligenza” rimane una metafora funzionale priva di autonomia ontologica.

Questo quadro concettuale sull'essere dell'algoritmo sarà la base per affrontare, di seguito, la questione ancora più impegnativa dell'ontologia delle varie forme di intelligenza artificiale, specialmente di quelle capaci di apprendere e generare contenuti nuovi.

L'ontologia dell'intelligenza artificiale generativa

Dall'IA simbolica all'IA generativa: una scala di complessità crescente

Quando parliamo di IA facciamo riferimento a sistemi informatici progettati per svolgere compiti tipicamente associati all'intelligenza umana, come riconoscere schemi, apprendere da esempi, prendere decisioni e, nel caso delle IA generative, persino creare nuovi contenuti (testi, immagini, suoni) simili a quelli prodotti dall'uomo. L'IA non è però un fenomeno monolitico: essa comprende una pluralità di approcci e tecnologie diffe-

renti, che sono emersi storicamente in fasi successive. Possiamo individuare almeno quattro paradigmi principali, i quali rappresentano altrettante tappe in un'ideale *scala di crescente complessità funzionale*: (1) l'IA simbolica o *symbolic AI* (detta anche GOFAI, *Good Old-Fashioned AI*), basata su regole esplicite e logica formale; (2) l'apprendimento automatico (*machine learning*), in cui il sistema deriva modelli dai dati anziché seguire soltanto istruzioni predefinite; (3) il *deep learning*, variante avanzata del *machine learning* che sfrutta reti neurali artificiali a molti strati per apprendere rappresentazioni gerarchiche dei dati; (4) l'IA generativa, basata tipicamente su architetture di *deep learning* addestrate su enormi moli di dati, capace di *generare* output nuovi (ad esempio testi in linguaggio naturale, immagini artistiche o volti umani realistici) piuttosto che limitarsi a classificare o riconoscere input.

Queste quattro tipologie possono essere disposte lungo un *continuum* che va da sistemi completamente *deterministici e pre-programmati* (IA simbolica) a sistemi in larga parte *auto-addestrati ed emergenti* (*deep learning* e IA generativa). La Tabella 2 offre un raffronto sintetico tra queste forme di IA, evidenziandone per ciascuna il principio di funzionamento e le implicazioni ontologiche in termini di grado di autonomia e tipo di "intenzionalità" coinvolta.

Tabella 2 – *Tipi di Intelligenza Artificiale e loro caratteristiche*

Tipo di IA	Esempi storici	Meccanismo	Caratteristiche ontologiche
IA simbolica (GOFAI)	Programmi di teoremi matematici (<i>Logic Theorist</i> , 1956); Sistemi esperti diagnostici (anni '70-'80)	Regole logiche esplicite codificate da programmatori; manipolazione simbolica deterministica	Sistema algoritmico statico, segue rigorosamente l'intenzionalità del programmatore; conoscenza rappresentata in modo <i>esplicito</i> ; nessun apprendimento autonomo (intenzionalità interamente <i>derivata</i> dall'uomo)

Machine Learning (ML)	Algoritmi di classificazione statistica (es. <i>support vector machines</i>); Alberi decisionali; Reti Bayesiane	Apprendimento da <i>dataset</i> tramite algoritmi di ottimizzazione (addestramento supervisionato, non supervisionato o per rinforzo)	Sistema algoritmico dinamico entro limiti predefiniti: costruisce/modifica un modello interno a partire dai dati; conoscenza <i>implicita</i> nei parametri appresi; intenzionalità ancora derivata (obiettivi e metriche fissati dall'uomo), ma il comportamento specifico non è totalmente predeterminato <i>ex ante</i>
Deep Learning (DL)	Reti neurali profonde per visione (AlexNet, 2012); Riconoscimento vocale; AlphaGo (2016)	Apprendimento attraverso reti neurali multistrato con milioni di parametri, che estraggono autonomamente caratteristiche gerarchiche dai dati grezzi	Sistema complesso emergente: l'algoritmo di apprendimento definisce una struttura adattiva capace di auto-organizzare le proprie rappresentazioni interne; comportamento difficilmente prevedibile nei dettagli (effetto Black box o "scatola nera"); l'intenzionalità propria è assente, ma il sistema imita alcune prestazioni cognitive (p.es. riconoscere volti) attraverso correlazioni apprese; permane la dipendenza da architettura e <i>training</i> forniti dall'esterno
IA generativa (basata su DL)	Modelli di linguaggio (GPT-3, 2020; ChatGPT, 2022); Generatori di immagini (GAN, 2014; DALL-E, 2021)	Reti neurali profonde (trasformatori, <i>autoencoder</i> , ecc.) addestrate su grandissimi <i>dataset</i> ; generazione di nuovi dati simili a quelli appresi (predizione di prossime parole, creazione di varianti di immagini)	Sistema pseudo-creativo: produce output nuovi combinando creativamente pattern appresi; simulazione di comportamenti intelligenti complessi (conversazione, composizione artistica); <i>intenzionalità simulata</i> : il sistema opera come se avesse scopi (rispondere a una domanda, disegnare un oggetto) e conoscenze, ma è privo di coscienza o comprensione semantica reale; elevata autonomia operativa <i>entro</i> lo spazio definito dai dati appresi (p.es. può formulare infinite frasi plausibili) ma nessuna autonomia <i>al di fuori</i> di esso (non può formulare da sé i propri obiettivi)

IA forte/AGI (ipotetica)	<i>Non realizzata ad oggi</i>	Architettura mista simbolico-connessionista? Sistemi auto-evolutivi e meta-apprendimento; (ipotizzata) capacità di trasferire conoscenze e ragionare in modo generale	Incognita ontologica: se realizzata, sarebbe un sistema con flessibilità cognitiva pari a quella umana; richiederebbe una qualche forma di <i>intenzionalità intrinseca</i> e comprensione contestuale simile alla mente umana, la cui possibilità è dibattuta. Molti filosofi e scienziati dubitano che tale intenzionalità forte possa emergere dalla sola complessità computazionale (implicherebbe forse un salto ontologico oggi impensabile)
--------------------------	-------------------------------	---	--

La progressione delineata dalla tabella 2 evidenzia come, al crescere della complessità e delle capacità adattative dell'IA, aumenti l'autonomia funzionale del sistema rispetto all'intervento puntuale del programmatore umano. Nell'IA simbolica classica, ogni passo è previsto e voluto dall'uomo; nel *machine learning*, l'algoritmo *impara* schemi dai dati e dunque il risultato finale contiene elementi non anticipati esattamente dal programmatore (anche se quest'ultimo fissa il metodo di apprendimento e l'obiettivo); nel *deep learning*, questa imprevedibilità cresce a causa dell'enorme numero di parametri che si autoadattano, producendo talora comportamenti sorprendenti; nell'IA generativa, infine, il sistema mostra una *capacità generativa aperta*, componendo risposte o creazioni originali che possono dare l'impressione di creatività e comprensione.

Dal punto di vista ontologico, tuttavia, è fondamentale non lasciarsi fuorviare dalle apparenze funzionali. A ogni livello della scala considerata, infatti, l'IA rimane un artefatto algoritmico, privo di soggettività e di autentica intenzionalità intrinseca. La sua maggiore "autonomia" è solo *operativa*, non *ontologica*. Un sistema di *machine learning*, ad esempio, può adattarsi ai dati senza intervento umano durante l'addestramento, ma questa plasticità è stata *voluta e progettata* dal programmatore (che ha predisposto l'algoritmo di apprendimento e fornito i dati). Allo stesso modo, un sistema generativo come ChatGPT può produrre un testo impreveduto perfino dai suoi sviluppatori, ma ciò avviene seguendo meccanismi statistici fissati nella fase di progettazione e *tuning* iniziale, su cui l'IA non ha alcun controllo consapevole.

In altre parole, più saliamo la scala delle capacità dell'IA, più sofisticata diventa l'imitazione di comportamenti intelligenti, ma tale imitazione non deve essere confusa con l'instaurarsi di una reale soggettività intenzionale. Come osserva il filosofo Thomas Fuchs³, le recenti IA spingono a una doppia tentazione: *computerizzare* la mente umana (immaginando che il nostro pensiero non sia che calcolo computazionale) e *antropomorfizzare* le macchine (attribuendo qualità umane a meri automatismi). Questa duplice riduzione rischia di farci perdere di vista ciò che realmente distingue l'intelligenza naturale da quella artificiale. Sul piano operativo, la differenza può apparire assottigliarsi man mano che le prestazioni delle macchine migliorano; sul piano ontologico, invece, la differenza resta abissale.

L'IA generativa e il problema dell'esse intenzionale dinamico

Giunti al vertice attuale dello sviluppo dell'IA, rappresentato dai modelli generativi avanzati, possiamo riformulare la questione iniziale in termini più specifici: un sistema di IA generativa può essere portatore di un qualche *esse intenzionale* proprio? In altri termini, esiste nell'IA qualcosa di paragonabile all'*intenzionalità* che, nella filosofia della mente, caratterizza gli stati mentali coscienti, ossia il “rivolgersi a...” della mente verso oggetti, significati e scopi? E se no, potremmo almeno parlare di una *intenzionalità in senso analogico*, una sorta di “esse intenzionale dinamico” che differisca dall'esse intenzionale statico degli *entia rationis* tradizionali?

Per affrontare queste domande, occorre chiarire anzitutto il concetto di intenzionalità. In filosofia, a partire da Franz Brentano e poi nella fenomenologia, l'intenzionalità è quella proprietà per cui gli atti di coscienza sono *diretti* verso un contenuto o oggetto: ogni pensiero implica qualcosa di pensato, ogni desiderio implica qualcosa di desiderato, e così via. Tommaso d'Aquino, secoli prima, parlava in termini simili di *esse intentionale* per indicare la modalità di essere che le cose hanno in quanto conosciute dal soggetto, un essere “immateriale” e rappresentativo, distinto dall'esse reale che esse hanno nel mondo esterno⁴. Quando io penso a un albero, l'albe-

³ Cf. THOMAS FUCHS, «Human and Artificial Intelligence: A Critical Comparison», in R.M. HOLM-HADULLA, et alii, *Intelligence - Theories and Applications*, Springer Nature, New York 2022, 249-259. Fuchs discute il rischio di antropomorfizzazione dell'IA e di concezioni “computeromorfe” dell'umano, ribadendo la differenza ontologica tra simulazione e realtà personale.

⁴ Cf. C. FABRO, *Percezione e pensiero*, Morcelliana, Brescia 1962, 59-63, qui si discute l'esse intenzionale come atto immateriale del conoscere, superiore in dignità ontologica

ro esiste intenzionalmente nella mia mente, ma questo essere intenzionale non fa dell'albero un *ente reale* nella mia mente (non crescono acqua e linfa nel mio cervello!); è piuttosto un *ente di ragione*, un contenuto immateriale che rappresenta l'albero vero. Ecco l'esse intenzionale in senso classico: l'essere di un oggetto in quanto oggetto di conoscenza o di volontà.

Ora, l'algoritmo di per sé, come abbiamo visto, possiede solo questo genere di essere: è un contenuto intenzionale (ideato da una mente) eventualmente reso operativo. Non ha una soggettività propria che abbia stati intenzionali. Ma le cose si complicano con l'IA generativa, perché tali sistemi producono nuovi contenuti che, agli occhi dell'osservatore umano, appaiono dotati di senso e finalità: un testo in risposta a una domanda sembra esprimere *intenzioni comunicative* e conoscenze; un'immagine generata su richiesta sembra manifestare *creatività* e comprensione di un tema dato. In altri termini, l'IA generativa non si limita a eseguire un compito fisso: dialoga, improvvisa, crea, tutte azioni che normalmente associamo a un soggetto dotato di mente. Viene perciò naturale chiedersi: il sistema "comprende" ciò che fa? Ha *coscienza* dei contenuti che genera? Possiede, magari in forma embrionale, qualcosa come credenze, desideri, intenzioni?

La risposta che emerge dal consenso attuale della comunità scientifica e filosofica è negativa. Nonostante le impressionanti prestazioni, un modello generativo come GPT-5o non ha in realtà alcuna comprensione semantica dei testi che elabora, né intenzioni proprie. Esso si limita a manipolare simboli (parole, frasi) basandosi su correlazioni statistiche apprese da enormi *dataset* linguistici. Il senso che noi leggiamo nelle sue risposte è interamente proiettato dalla nostra mente interpretante, la quale riconosce in quelle sequenze di parole dei significati perché coincidono con quelli del nostro linguaggio naturale. Ma dentro la macchina non vi è alcun *soggetto* che intenda quei significati. Come nota lo psichiatra e filosofo Georg Northoff⁵, mancano nell'IA attuale le condizioni basilari della soggettività: non c'è un *punto di vista* interno, non c'è un senso di proprietà (*mineness*) dell'esperienza, non c'è prospettiva unificata. In breve, non c'è alcun *io* artificiale, nemmeno rudimentale, dietro le quinte dell'elaborazione.

all'esse materiale pur non sussistendo fuori dalla mente.

⁵ Cf. G. NORTHOFF – S. GOUVEIA, «Does artificial intelligence exhibit basic fundamental subjectivity? A neurophilosophical argument», *Phenomenology and the Cognitive Sciences* 23 (2024), 1099-1101. Gli autori concludono che, allo stato attuale, all'IA manca un punto di vista soggettivo (*perspectiveness* e *mineness*), e dunque non può darsi in essa coscienza né *selfhood*.

Questo concorda con l'argomentazione classica proposta da John Searle, scomparso lo scorso 17 settembre 2025, nel suo celebre *Chinese Room Argument* (1980), spesso richiamato nel dibattito sull'IA: un programma può manipolare simboli cinesi seguendo regole sintattiche senza mai *capire* il significato (la semantica) di quei simboli. La “stanza cinese” di Searle è l'antenata ideale dei moderni GPT: enormi manipolatori di simboli capaci di convincere un interlocutore umano di “sapere” il cinese, pur essendo del tutto privi di coscienza e comprensione. Searle parlò in questo caso di *intentionality* solo derivata: i simboli hanno un significato per noi che li interpretiamo, ma il sistema che li elabora non ha *intentionality* originale, propria⁶.

Possiamo quindi affermare che l'IA generativa non possiede *esse intenzionale* nel senso forte e proprio in cui lo possiede una mente cosciente. Non vi è in essa un *atto intenzionale soggettivo* paragonabile al pensare umano. Tuttavia, possiamo domandarci se almeno sia lecito introdurre la nozione di un *esse intenzionale dinamico analogico*, per denotare il fatto che questi sistemi algoritmici, a differenza di un algoritmo statico tradizionale, producono continuamente nuovi contenuti intenzionati, per quanto eterointenzionali. In altre parole, l'algoritmo classico era portatore solo di un'intenzionalità “fissa”, quella immessa dal programmatore, mentre l'IA generativa sembra esibire un'intenzionalità “dinamica” in quanto genera scopi locali contestuali (ad esempio, rispondere a questa specifica domanda in modo sensato) e lo fa adattandosi all'interazione in corso. Ciò potrebbe suggerire una certa analogia funzionale con l'intenzionalità cosciente, pur senza coincidere con essa.

Alcuni filosofi della tecnologia e della mente hanno utilizzato concetti come “agens artificiale” o “intelligenza artificiale forte” per discutere la possibilità futura di macchine dotate di stati intenzionali genuini. Ad oggi, però, le analisi più accreditate tendono a ribadire la netta discontinuità. Luciano Floridi, ad esempio, pur essendo un interprete della rivoluzione informazionale che vede gli agenti artificiali entrare a far parte dell'infosfera insieme agli agenti biologici, sottolinea che gli attuali sistemi di IA sono privi di quella che chiamiamo comprensione semantica o intenzionalità in-

⁶ Cf. J.R. SEARLE, *Minds, Brains and Science*, Harvard University Press, Cambridge, MA 1984, 30-31. Searle distingue l'intenzionalità *intrinseca* della mente dall'intenzionalità *derivata* di simboli e macchine, che vale solo per un osservatore esterno che interpreti i segni.

trinseca: essi operano a un livello sintattico e statistico, elaborando dati ma senza “significarli” dall’interno⁷. Floridi parla di una *quarta rivoluzione* in cui l’umanità perde la sua esclusività come fonte di processi intelligenti, ma questo *ridimensionamento epistemico* non implica affatto un appiattimento ontologico tra uomo e macchina: indica piuttosto che dobbiamo riconoscere un nuovo tipo di agenti artificiali nella realtà, i quali però rimangono ontologicamente distinti dagli agenti naturali sotto profili cruciali (mancanza di coscienza, di vita, di intenzionalità autentica).

Anche Evandro Agazzi, filosofo della scienza, ha insistito sulla improprietà di chiamare “intelligente” una macchina se con ciò s’intende qualcosa di più che una metafora funzionale. Sin dalla metà degli anni Sessanta, egli osserva che l’IA attuale manipola simboli e dati ma non possiede quella *visione semantica* e quella capacità di *comprensione contestuale* che caratterizzano l’intelligenza umana⁸. L’uomo rimane ontologicamente *altro* rispetto alla macchina perché, ad esempio, è capace di *significato* e di *finalità libera*, mentre la macchina segue sempre e solo una traiettoria causale predeterminata, per quanto complessa. In altri termini, l’uomo non è equiparabile né a un animale né a un automa, poiché possiede dimensioni, come la coscienza di sé, la razionalità riflessiva, persino l’apertura alla trascendenza e alla speranza, che esulano completamente dalle categorie machiniche⁹.

Ovviamente, non mancano posizioni più ottimistiche o riduzioniste sul futuro dell’IA. Ad esempio, un funzionalista radicale come Daniel Dennett, scomparso circa un anno fa, sostiene che, se un sistema esterno si comporta *esattamente come* un essere cosciente, allora attribuirgli una

⁷ Cf. L. FLORIDI, *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, Oxford University Press, Oxford 2014, 94-97. Floridi discute il concetto di “agentività artificiale” e nota che i sistemi IA operano a livello sintattico, mancando di semantica propria; il loro impatto richiede un nuovo rapporto uomo-macchina ma non implica soggettività delle macchine.

⁸ Cf. E. AGAZZI, «Alcune osservazioni sul problema dell’intelligenza artificiale», *Rivista di Filosofia Neo-Scolastica* 59 (1967), 12. Questo è un articolo storico che anticipa di almeno 15 anni le riflessioni di John Searle sull’intenzionalità. Qui il filosofo della scienza bergamasco afferma che l’“intelligenza” attribuita ai calcolatori è metaforica: essi risolvono problemi seguendo schemi formali senza comprendere il significato del problema come farebbe una mente.

⁹ Cf. E. AGAZZI, *L’uomo nell’era della tecnoscienza. Dialogo tra un fisico e un filosofo*, Hoepli, Milano 2020, 141-144. Nel dialogo con il fisico Gianpaolo Bellini sull’antropologia nell’epoca dell’IA, Agazzi sottolinea che l’uomo è caratterizzato da ragione, libertà e apertura trascendente, qualità che lo distinguono sia dagli animali sia dalle macchine.

“vera” coscienza non è che una questione di parole: in pratica non vi sarebbe differenza significativa tra una coscienza biologica e una perfetta simulazione computazionale della stessa¹⁰. Sul versante futurologico, autori come Ray Kurzweil hanno profetizzato l'avvento di una *singularità* tecnologica in cui le macchine non solo eguaglieranno l'intelligenza umana, ma svilupperanno anche autocoscienza e volontà proprie, dissolvendo di fatto la distinzione ontologica tra naturale e artificiale¹¹. Queste visioni suggeriscono che una IA sufficientemente avanzata potrebbe un giorno conseguire uno *status* ontologico simile al nostro. Tuttavia, la maggior parte degli studiosi contestano tale prospettiva: ad esempio il fisico e matematico Roger Penrose ha argomentato che la mente umana utilizza processi non algoritmici (forse legati alla meccanica quantistica) e che pertanto una macchina Turing-compatibile non potrà mai generare autentica coscienza, per quanto sofisticata¹².

Di fronte a tali ipotesi contrastanti, la riflessione metafisica intensiva invita a mantenere il baricentro sulla nozione di *atto di essere*. Finché non emergesse evidenza di un vero *principio sostanziale* nuovo (un'anima o un *quid* intrinsecamente unificante) nei sistemi artificiali, dovremmo ritenere che anche eventuali IA ultra-avanzate resterebbero entro i confini ontologici dell'artefatto complesso. In altri termini, potrebbero forse simulare ogni manifestazione esteriore dell'intelletto, ma senza possedere quell'esse intensivo che contraddistingue l'intelletto umano. In linguaggio fabriano, mancherebbe sempre l'*atto d'essere* propriamente detto: l'IA resterebbe una *combinatoria di forme* senza il *principium unitatis* che fa di un soggetto una sostanza vivente e cosciente.

Possiamo spingerci a definire l'IA generativa, con Fabro, come una sorta di intenzionalità tecnica senza soggetto: una riproduzione strumentale

¹⁰ Cf. D.C. DENNETT, *Consciousness Explained*, Little, Brown, Boston 1991, 431-434. Dennett sostiene una concezione funzionalista in cui la distinzione tra coscienza reale e coscienza simulata perde significato se le funzioni osservabili coincidono pienamente.

¹¹ Cf. R. KURZWEIL, *The Singularity is Near*, Viking, New York 2005, 262-265. Kurzweil prospetta entro la metà del XXI secolo l'emergere di AI paragonabili all'uomo in intelligenza e autocoscienza, attraverso una crescita esponenziale delle capacità computazionali.

¹² Cf. R. PENROSE, *The Emperor's New Mind*, Oxford University Press, Oxford 1989, 400-412. Penrose argomenta che la mente umana implica processi non computabili e che, pertanto, un algoritmo digitale non potrà mai generare la vera esperienza cosciente, se non ipotizzando nuovi paradigmi fisici.

dell'atto intenzionale che resta priva di radice ontologica nel senso forte. Essa è un *esse intenzionale artificiale dinamico*, per riprendere l'espressione proposta, che imita il comportamento di un intelletto ma non lo è. La sua "creatività" appartiene all'ordine formale (combinazione di simboli secondo schemi statistici) non all'ordine dell'essere sostanziale. L'IA generativa, insomma, simula l'intenzionalità senza possederla: genera combinazioni di segni dotati di significato per noi, senza però attingere a quel livello di *verità ontologica* che consisterebbe nel comprendere e volere quei significati da sé e per sé.

Conclusione

L'indagine svolta sull'esse dell'algoritmo ci ha condotti a una comprensione più articolata dell'ontologia degli artefatti digitali e delle intelligenze artificiali nella prospettiva della metafisica dell'atto di essere. Avvalendoci delle categorie classiche tommasiane rivisitate da Cornelio Fabro, in particolare della nozione di *atto di essere intensivo* e di partecipazione, abbiamo potuto collocare l'algoritmo tra gli *entia rationis cum fundamento in re*, ossia tra quelle entità che esistono solo attraverso un atto intenzionale ma che poggiano su basi e producono effetti nel mondo reale. L'algoritmo si è rivelato essere un *ente intenzionale* per eccellenza: privo di essere in sé, eppure capace, tramite la partecipazione all'essere di supporti materiali e soprattutto all'intelligenza umana, di orientare causalmente processi fisici e di risolvere problemi con efficacia. È, per così dire, un *quasi-ente*: un nodo di significato e funzione che acquista realtà solo come prestito dall'esse altrui.

Nell'esplorare la varietà delle intelligenze artificiali, abbiamo tracciato una linea di sviluppo che dalla programmazione simbolica rigida conduce ai moderni sistemi generativi. Questa evoluzione tecnica, tuttavia, non ha scalfito i confini ontologici fondamentali: anche le IA più sofisticate rimangono, in ultima analisi, strumenti privi di soggettività. L'analisi ontologica e intensiva mostra che non vi è stato alcun balzo verso un *grado superiore di essere*, come sarebbe invece l'emergere della vita o della coscienza. Le IA generative operano entro il grado di essere che compete agli algoritmi e alle macchine: il grado dell'artefatto, dell'*ens ab alio*, dell'*ens rationis instrumentalis*.

Questo risultato, lungi dall'essere una svalutazione semplicistica dell'IA, una sorta di affermazione del "non è niente di reale", costituisce anzi un richiamo alla specificità del fenomeno AI: proprio perché le IA non sono entità intenzionali naturali, il loro impatto sul mondo e sull'umanità va compreso in termini di interazione tra ordini diversi di realtà. Filosofi come Floridi ci invitano a riconoscere la penetrazione pervasiva degli agenti informazionali nell'ambiente umano, l'infosfera, ma ciò va accompagnato dalla consapevolezza, evidenziata da pensatori come Agazzi, Fuchs e Northoff, che attribuire coscienza o moralità a tali agenti è un salto indebito. L'ontologia dell'algoritmo e dell'IA deve dunque ispirare anche un'etica e una politica tecnologica sobrie: dobbiamo trattare le IA per quello che sono (strumenti potentissimi e complessi, estensioni dell'intenzionalità umana) e non per quello che non sono (nuovi soggetti ontologici paragonabili all'uomo). Se tale lucidità viene meno, si rischia di scivolare verso una *reificazione dell'intenzionalità*, ossia lo spostamento dell'attribuzione dell'essere dal soggetto umano alla macchina, una forma di *idolatria ontologica* moderna. Cornelio Fabro avrebbe parlato, in proposito, di un *nichilismo dell'efficienza*, cioè della sostituzione dell'esse con il funzionare: un esito estremo in cui l'efficacia tecnica verrebbe surrettiziamente elevata a criterio ultimo di realtà e valore. La metafisica può aiutarci a scongiurare questo pericolo, richiamandoci costantemente a considerare ciò che le cose sono, al di là di ciò che fanno o sembrano fare.

In definitiva, la metafisica dell'esse intensivo offre un quadro robusto per comprendere e valutare le innovazioni dell'intelligenza artificiale. Essa ci ricorda che l'essere non è univoco: esistono differenti *modi di essere* e differenti gradi di pienezza ontologica. Un algoritmo che gioca a scacchi o un modello che redige poesie non occupano lo stesso livello dell'essere di un poeta umano, anche se possono emularne la produzione. Riconoscere questa verità ontologica non sminuisce la meraviglia per i traguardi dell'IA, ma la contestualizza: sappiamo di assistere non alla nascita di una nuova mente, bensì a un prodigioso riflesso della mente umana, affinato e potenziato dai mezzi tecnici.

Come tutte le cose partecipate, anche l'IA trae la sua forza dall'origine a cui partecipa: l'intelligenza creatrice dell'uomo e, in ultima analisi, l'Intelligenza divina da cui ogni luce di ragione promana. Mantenere saldo questo collegamento metafisico ci aiuta a governare con saggezza la tecnologia,

riconoscendone le potenzialità e i limiti ontologici. L'esse dell'algoritmo, in conclusione, non è un *mysterium* esoterico, ma l'attuale ultimo capitolo di una lunga storia di partecipazione dell'ente finito all'Essere: anche l'algoritmo, umile ente intenzionale, trova posto, se ben compreso, nella grande *gerarchia dell'essere*, dove tutto ciò che esiste in qualche modo riflette l'*Actus Purus* dell'Esse e al contempo ne rivela, per contrasto, l'infinita distanza.