

FONDAZIONE CENTESIMUS ANNUS

CONFERENZA INTERNAZIONALE

“INTELLIGENZA ARTIFICIALE GENERATIVA”.

Intervento del Cardinale Pietro Parolin, Segretario di Stato

Sala Clementina

Sabato 22 giugno, ore 11:30

Signora Presidente,
Cari amici,

rivolgo a tutti voi un cordiale saluto e un sentito ringraziamento per i lavori svolti in occasione di questa Conferenza Internazionale promossa dalla Fondazione “Centesimus Annus pro Pontefice”. Il vostro impegno nell’approfondimento della Dottrina Sociale della Chiesa è un prezioso contributo alla missione di annuncio del Vangelo di tutta la Chiesa. L’ascolto della Parola di Dio chiede una costante attitudine all’ascolto delle domande e degli interrogativi di senso dell’uomo di oggi, rendendo ragione della fede nel confronto con una realtà in continuo movimento.

Tra le sfide che si sollevano dall’attualità si pone certamente quella rappresentata dall’evoluzione delle tecnologie digitali basate sull’Intelligenza Artificiale, la cui rapida diffusione si insinua in maniera pervasiva nella quotidianità, influenzando abitudini individuali e comportamenti sociali.

Il loro impatto sulla globalizzazione dell’economia, dei mercati del lavoro, sugli indirizzi della politica e le espressioni culturali appare sempre più evidente, tanto da far supporre che ci confrontiamo non soltanto con un’epoca di cambiamenti, ma con vero e proprio cambiamento d’epoca.

A subire una profonda modificazione è il modo in cui l’uomo comprende se stesso, nel leggere il presente e immaginare il futuro, generando una narrazione del suo essere gettato nel mondo che avanza la pretesa di una discontinuità con la grammatica della vita ereditata dalla tradizione umanistica.

Rispetto alle grandi innovazioni tecnologiche del passato – come ad esempio l’invenzione della stampa, della macchina a vapore o dell’automobile – la “rivoluzione algoritmica” sembra imporre un radicale cambio di paradigma: mentre gli artefatti costruiti in precedenza dall’uomo sono serviti per trasformare il mondo fisico, le nuove tecnologie informatiche processano una realtà immateriale, ovvero l’informazione. Dal momento che intervengono nei processi di produzione, immagazzinamento e gestione della conoscenza, possono esercitare un significativo impatto sul modo in cui si sviluppa l’intelligenza umana. Esternalizzando funzioni cognitive, come la memoria, ed estendendo capacità mentali, come la percezione, rendono più labile il confine tra natura e cultura, dando luogo ad un vero e proprio “ambiente digitale”.

Una «*trasformazione strutturale nell’esperienza dell’io e dei suoi rapporti con gli altri*»¹ che Romano Guardini aveva già segnalato nei primi decenni del XX secolo, presagendo come il connubio tra tecnica e potenza avrebbe comportato una complessificazione degli interrogativi etici. L’uomo di oggi, notava con lucidità il teologo, si trova nella condizione di essere soggetto dell’innovazione tecnologica e allo stesso tempo oggetto delle sue imprevedibili conseguenze: «*ha potere sulle cose, ma non ha ancora potere sul proprio potere*»².

Una sorta di strano paradosso che si pone all’origine di un «*sovraccarico esistenziale*»³, di una diffusa e indefinita paura nei confronti dell’avvenire: se da una parte aumenta la fiducia nella tecnologia, tanto da attribuirle un valore quasi salvifico e «*supporre che ogni problema futuro possa essere risolto con nuovi interventi tecnici*»⁴, dall’altra, aumenta anche il timore che le macchine possano sfuggano al controllo dell’uomo.

Non si vuole disconoscere gli importanti traguardi raggiunti dall’umanità, né negare gli effetti positivi del progresso tecnologico, quanto piuttosto richiamare all’urgenza di vigilare sul suo avanzamento, per non cedere a quella logica che si pone in modo acritico di fronte ad esso e che papa Francesco stigmatizza come «*pragmatismo fatale*»⁵.

¹ R. GUARDINI, *La fine dell’epoca moderna*, Morcelliana, Brescia 1993, 62.

² *Ibid.*, 87.

³ R. GUARDINI, «Europa – Realtà e compito», in ID. *Ansia per l’uomo*, vol. 1, Morcelliana, Brescia 1969, 281-282.

⁴ FRANCESCO, *Esortazione Apostolica Laudate Deum*, Roma, 4 Ottobre 2023, n. 57.

⁵ FRANCESCO, *Laudate Deum*, 57.

Comprendere le trasformazioni che già si stanno verificando sotto ai nostri occhi, a seguito del loro turbinoso dispiegarsi, è un atto necessario se vogliamo indirizzare la rivoluzione introdotta dalle IA in modo equo dal punto di vista sociale e sostenibile sotto il profilo ambientale.

È ciò su cui si interroga il Messaggio per la Giornata Mondiale della Pace 2024, in cui papa Francesco esorta ad acquisire più strumenti – sia individuale che collettivi – per gestire l’influenza che già oggi le IA esercitano su di noi. Avviare un esercizio congiunto di riflessione e discernimento sulle conseguenze etiche di questa nuova frontiera dell’agire umano, ha come obiettivo precipuo quello di costruire un futuro di giustizia e di pace per le generazioni a venire.

1. Intelligenze Artificiali: lo scenario attuale, tra ricerca e sperimentazione

Ad oggi, non esiste una definizione univoca di Intelligenza Artificiale (AI), sebbene il termine sia ormai entrato a far parte del linguaggio di uso comune. Nella sua accezione più ampia, con Intelligenza Artificiale si indica l’insieme di teorie e tecniche pratiche volte allo sviluppo di sistemi informatici in grado di imitare le funzioni cognitive superiori dell’uomo: apprendimento, memoria, deduzione, inferenza.

Attualmente, l’IA è oggetto di una fase di rapida evoluzione che segna il passaggio dall’apprendimento automatico (*Machine learning*), in cui risultava ancora fondamentale l’ausilio della programmazione e dell’istruzione da parte degli sperimentatori, a quello che viene definito apprendimento profondo (*Deep learning*), in cui si evidenzia la capacità degli algoritmi di acquisire i dati in modo autonomo, ottimizzando progressivamente le proprie prestazioni.

Vista l’ampia gamma di modelli, sperimentazioni e obiettivi in cui l’IA si declina, specifica e differenzia, è preferibile parlarne al plurale, così da sottolineare al contempo «*il divario incolmabile che esiste tra questi sistemi, per quanto sorprendenti e potenti, e la persona umana*»⁶.

⁶ *Ibid.*

Una breve panoramica sulla diversità di modellizzazioni delle IA può essere utile a prendere atto di come esse costituiscano solo un riflesso e un'espressione frammentaria dell'intelligenza umana.

È possibile, infatti, individuare quattro sottoinsiemi di IA: predittiva; basata sulle anomalie; decisionale; generativa.

Si parla di *IA predittiva* quando una macchina è in grado di avanzare previsioni sulla base di una combinazione di input acquisiti in precedenza e di un'analisi degli scenari attuali. Mediante algoritmi statistici e linee attive di *machine learning*, l'IA predittiva analizza i dati storici, al fine di identificare la ricorrenza di modelli o tendenze che possono essere applicati su ulteriori dati ancora non analizzati. Questa tipologia di IA è già ampiamente impiegata nel mercato azionario, nella diagnostica medica, nella previsione del comportamento dei consumatori. Tuttavia, il fatto di basare le previsioni unicamente sull'assunto secondo cui i modelli futuri seguiranno obbligatoriamente quelli passati, ne costituisce un limite evidente. Se ci si trovasse a fare i conti con un'inaspettata congiuntura di eventi o con un cambiamento improvviso nei dati, qualunque tipo di previsione risulterebbe inattendibile.

L'*IA basata sulle anomalie* è addestrata a riconoscere le regolarità e a rilevare le difformità che si presentano in uno schema. Risulta particolarmente utile per la sicurezza, poiché i metodi tradizionali di difesa informatica, che si basano sull'utilizzo di regole e di pattern predefiniti, si stanno dimostrando insufficienti nei confronti delle minacce dinamiche. Tramite analisi avanzate e tecniche di rilevamento, i sistemi di AI possono individuare le attività dannose, adattarsi ad esse in tempo reale e fornire strumenti di difesa proattiva.

L'*IA decisionale* estrae informazioni da grandi dataset, classificandole sulla base delle caratteristiche indicate dai programmatori. Collegando i dati elaborati ai risultati reali, fornisce suggerimenti volti ad indirizzare le scelte e offre riscontri utili a verificarne l'efficacia. L'aumento del volume delle decisioni, tuttavia, sta ponendo di fronte alla necessità di affidare alle macchine un ruolo esecutivo più determinante. L'aspetto quantitativo è essenziale per

comprendere la scala e l'ineluttabilità dell'autonomia decisionale algoritmica: se immaginassimo di distribuire uniformemente su ogni abitante del pianeta la totalità dei dati prodotti su base annua, ci accorgeremmo che siamo già nell'ordine di svariate decine di gigabyte pro-capite al giorno⁷.

Tuttavia, l'automatizzazione decisionale implica la significativa riduzione del coinvolgimento degli esseri umani e la conseguente delega alle macchine dell'esecuzione di compiti svolti in autonomia. Questo fatto pone un problema di *accountability*, cioè rende difficile determinare su chi ricadrebbe la responsabilità di dover rendere conto di eventuali danni.

Nel discorso che ha recentemente tenuto in occasione del G7, il Santo Padre ha manifestato la sua preoccupazione per questo tipo di IA, affermando che esiste una differenza fondamentale tra le scelte algoritmiche e la capacità umana di decidere: «*La decisione è un elemento che potremmo definire maggiormente strategico di una scelta e richiede una valutazione pratica. A volte, spesso nel difficile compito del governare, siamo chiamati a decidere con conseguenze anche su molte persone*»⁸. La decisione umana rimanda all'esercizio della sapienza, perché ingloba ragioni che possono rimanere escluse dal calcolo automatizzato, come l'aspirazione al bene comune e l'affermazione dell'inalienabile valore di ogni vita umana. Discernere e giudicare sono atti umani che trascendono il mero processo di *decision making* operato da una macchina.

Infine, l'IA generativa utilizza tecniche avanzate di *machine learning*, come le reti neurali, allo scopo di identificare all'interno di un dataset gli schemi e le relazioni sottostanti, così da generare output simili per stile e struttura ai grandi set dei dati di input. In poche parole, le reti neurali sono allenate a produrre nuovi esempi che imitino gli input di addestramento. E' però doveroso premettere che i risultati da esse elaborati non sono effettivamente "originali", in quanto l'IA generativa «*cerca nei big data delle informazioni e le confeziona nello stile che le è stato richiesto*»⁹. E' su questo punto che

⁷ <https://www.statista.com/statistics/871513/worldwide-data-created/> ↑

⁸ FRANCESCO, *Messaggio di Sua Santità papa Francesco al G7*, 14 Giugno 2024.

⁹ *Ibid.*

concentreremo ora la nostra attenzione per comprendere meglio quali sono i rischi e le opportunità cui esse aprono.

2. IA Generativa: rischi e opportunità di una nuova frontiera dell'agire umano

L'ampia gamma di possibilità applicative a cui le AI si aprono, ne mostra un potenziale ancora tutto da esplorare, lasciando intuire come possano giocare un ruolo cruciale nel miglioramento delle condizioni di vita dei singoli, così come di intere Nazioni. Digni di nota sono i risultati già ottenuti in campi come quello della logistica, del settore dei trasporti, della comunicazione, dell'economia, della didattica ecc. In ambito farmaceutico, la collaborazione tra intelligenza umana e artificiale ha consentito scoperte che gli scienziati non sarebbero stati in grado di raggiungere facendo affidamento sulle loro sole competenze. Gli sperimentatori del MIT di Boston, ad esempio, sono stati in grado di sintetizzare un nuovo antibiotico - l'hacilina - capace di contrastare ceppi di batteri fino a quel momento resistenti a qualsiasi altro farmaco esistente.

Nel Messaggio per la Giornata Mondiale della Pace 2024 si accenna a come l'impiego delle AI, se fosse orientato allo sviluppo umano integrale, «potrebbe introdurre importanti innovazioni nell'agricoltura»¹⁰. Attraverso l'uso di dispositivi da remoto, il monitoraggio delle colture destinate all'alimentazione potrebbe consentire l'ottimizzazione dei raccolti, permettendo ad intere popolazioni di incrementare il prodotto interno e, così, contrastare il problema della fame.

Allo stesso tempo, però, le IA si espongono ad un'ambivalenza che chiede di essere posta al vaglio, perché se consideriamo l'utilizzo che potrebbe farsi di queste tecnologie in altri ambiti, come ad esempio il settore degli armamenti, la possibilità di condurre azioni militari attraverso sistemi di controllo da remoto potrebbe avere un impatto devastante sull'umanità. Ciò renderebbe

¹⁰ FRANCESCO, *Messaggio per la LVII Giornata Mondiale per la Pace*, 1 Gennaio 2024, n. 6.

più arduo contenere i conflitti e assicurare relazioni di sicurezza tra Paesi, acuendo il rischio di escalation e perdite di vita umana.

Con il crescente impiego dell'IA, aumenta la preoccupazione sulla privacy, dal momento che i modelli generativi possono inavvertitamente acquisire informazioni sensibili o personali dai dati di allenamento (*data breach*). Il basso costo di archiviazione implica che i dati possano esistere più a lungo di quanto necessario ed essere riutilizzati per finalità diverse da quelle previste.

Un problema serio è dato anche dall'utilizzo dell'IA generativa con l'intento di creare contenuti illegali, come *e-mail di phishing*, documenti contraffatti, furto d'identità. Un'ulteriore criticità è rappresentata dalla possibilità che le IA generino contenuti che violino le proprietà intellettuali (copyright, marchi o brevetti).

Le opportunità, ma anche i rischi, derivanti dalla diffusione di questi dispositivi, sono dunque di portata tale da aver indotto l'Unione Europea a mettere a punto una specifica regolamentazione in materia. L'*AI Act* istituisce un quadro giuridico uniforme che si pone l'obiettivo di normare lo sviluppo, la commercializzazione e l'uso dei sistemi di AI, basandosi su una classifica di quattro livelli di rischio (inaccettabile, elevato, limitato, minimo o nullo).

Le aziende che si situano ad un livello di rischio limitato avranno solo l'obbligo di informare gli utenti circa il funzionamento e l'utilizzo del loro prodotto, mentre per i servizi classificati a rischio elevato, in cui rientrano quelli offerti dalla gran parte di sistemi di IA generativa, sono previsti divieti più stringenti. In particolare, i contenuti generati da sistemi cibernetici che alterano la realtà - foto, video, audio - dovranno essere contrassegnati da un'etichetta che consenta ai fruitori di discernere la loro derivazione. Sono vietate, invece, tutte le tipologie di IA classificate a rischio inaccettabile, come gli strumenti di polizia predittiva, i sistemi di riconoscimento biometrico, di decodifica delle emozioni, di *social scoring*. Occorre tenere presente, però, che la gran parte delle aziende che sviluppano tecnologie basate sull'IA generativa non ha sede in Europa, per cui non è soggetta alle norme previste dall'*AI Act*. Ciò rende necessario e urgente il raggiungimento di un accordo internazionale,

quanto meno sotto la forma di un codice di condotta volontario, che stabilisca dei parametri uniformi per le aziende del settore.

Dal punto di vista delle ricadute sociali, il guanto di sfida lanciato dall'IA si pone sul fronte del lavoro e della custodia dell'ambiente.

Una delle trasformazioni più allarmanti, registrate sia a livello europeo che internazionale, è la polarizzazione della forza lavoro. Si tratta di una tendenza che segnala l'aumento delle richieste di occupazioni nelle professioni ad elevata specializzazione a cui, parallelamente, segue una diminuzione in quelle a bassa qualifica. Si prevede che le IA affiancheranno l'uomo nei lavori composti da molti compiti complessi, con un cospicuo aumento della produttività. Ciò implicherà che, in attività a basso livello di competenza, le IA forniranno al lavoratore umano indicazioni sullo svolgimento di alcuni compiti specifici, semplificando le procedure, ma allo stesso tempo rendendo i processi meno trasparenti. Riducendo il livello di addestramento dei lavoratori umani, le aziende potrebbero sostituire più facilmente i propri dipendenti e, di conseguenza, abbattere i costi di produzione con l'abbassamento dei salari. Per contrastare il disallineamento delle competenze (*skill-mismatch*) dei lavoratori, è necessario progettare politiche che reimmettano nel tessuto sociale valori come la solidarietà, investendo sulla formazione, in modo che i dipendenti maggiormente colpiti dall'AI possano effettuare una transizione professionale efficace. Allo stesso tempo, occorre anche sostenere l'imprenditoria, incoraggiando le nuove forme di mutualismo rese possibili dagli strumenti digitali, come le piattaforme di cooperazione.

Per quanto concerne la tutela e la cura della "Casa comune", invece, il successo delle IA potrebbe comportare un altissimo prezzo in termini di impatto sulle risorse rinnovabili. L'estrattivismo minerario di litio, rame, coltan ed altre terre rare desta grande preoccupazione per la minaccia rappresentata alla biodiversità e alla sopravvivenza delle comunità residenti. Le conseguenze dovute non solo alle fasi di localizzazione dei filoni e di perforazione del suolo, ma anche di trasporto, produzione, lavorazione delle materie prime, può avere effetti irreversibili sull'ambiente. L'impiego di metalli pesanti e tossici come agenti di fusioni pone il problema dello

smaltimento dei rifiuti, spesso riversati nelle falde acquifere, con conseguente contaminazione di fiumi e torrenti.

La protezione dell'acqua è al centro anche delle questioni direttamente legate all'incremento delle tecnologie digitali. Infatti, l'addestramento e l'esecuzione di modelli AI su larga scala non soltanto comporta un ingente dispendio di energia, con conseguente aumento delle emissioni di CO₂, ma anche un consumo di acqua che potrebbe non essere sostenibile a medio e lungo termine. Per fare un esempio concreto, i ricercatori hanno calcolato che l'addestramento di ChatGpt-3 ha consumato settecentomila litri di acqua dolce per il raffreddamento dei moderni centri di elaborazione dati di Microsoft.

Occorre rendere obbligatorio il tracciamento dell'impronta idrica dei colossi del settore informatico, dal momento che insistere sulla trasparenza è il primo passo verso un uso responsabile delle risorse e la pianificazione di obiettivi di sostenibilità, tra cui dovrebbe porsi come prioritario il reintegro entro il 2030 di una quantità d'acqua superiore a quella finora utilizzata.

3. Conclusione

L'immensa crescita tecnologica necessita di un accompagnamento, di un'adeguata educazione alla responsabilità circa le finalità dei suoi futuri sviluppi. Nessuna tecnologia è per se stessa "neutra", nel senso di essere culturalmente indifferente, disincarnata, ma è un prodotto pienamente umano, i cui orientamenti si compongono di scelte che, di epoca in epoca, sono condizionati da valori individuali, sociali e culturali.

Papa Francesco ha indicato due criteri fondamentali e irrinunciabili nella valutazione delle nuove tecnologie: il rispetto della dignità di ogni essere umano e la fraternità tra tutti gli uomini e le donne. Afferma con convinzione che: «uno sviluppo tecnologico che non comportasse un miglioramento della qualità di vita di tutto il genere umano, non potrebbe essere considerato un reale avanzamento»¹¹. Se fosse altrimenti, se le tecnologie digitali fossero

¹¹ FRANCESCO, *Messaggio di Sua Santità papa Francesco per la LVII Giornata Mondiale per la Pace*, n.2.

appannaggio di una piccola parte di umanità, le disuguaglianze potrebbero crescere a dismisura: non solo le ricchezze, ma anche il potere che deriva dal possesso di queste nuove forme di sapere, potrebbe concentrarsi nelle mani di pochi. Le IA diverrebbero un ulteriore e poderoso strumento che andrebbe a rafforzare e consolidare i dinamismi già in atto nel paradigma tecnocratico. In tal senso, è importante chiedersi quali forme di potere l'IA semplifichi, riproduca o abiliti, quali interessi promuova e chi ne sia responsabile (*accountability*).

All'approccio dell'algoritica, che vede le intelligenze artificiali come agente morale e che delinea un'etica dell'intenzione, è necessario affiancare anche un'etica della conseguenza, cioè che valuti la moralità di un'azione dal risultato che produce. Per migliorare gli algoritmi è necessario sviluppare una cornice sociotecnica, in cui i progetti siano svolti con modalità che non solo promuovono obiettivi benefici, ma che si orientino a questo risultato con modalità socialmente eque e sostenibili.

Questa è la sfida che si delinea di fronte a noi e che interpella in modo speciale i credenti in Cristo: scorgere in quest'epoca di trasformazioni, un tempo propizio e opportuno a riorientare il cammino dell'umanità verso il Signore e verso gli altri, impegnandoci a riscoprire la nostra comune appartenenza alla Casa comune, nella tensione verso la costruzione di un mondo più giusto e fraterno. Sarebbe riduttivo, però, ricondurre questo cambio di passo alla sola dimensione tecnico-economica, perché gli interventi legislativi e le decisioni politiche – per quanto necessari – non sono sufficienti ad assicurare che queste nuove tecnologie siano poste al servizio della fraternità e della pace.

Di fronte a una rivoluzione che è già entrata nelle nostre case ed è destinata a cambiare in modo profondo i nostri stili di vita, il modo in cui ci informiamo, impariamo, comunichiamo o scegliamo di vivere il nostro tempo libero, è necessario un importante investimento dal punto di vista formativo, rivolto a tutte le età, per comprendere in che modo funzionano queste macchine, che cosa possiamo aspettarci da loro e su che cosa dobbiamo criticamente vigilare.

Abbiamo bisogno di un'alfabetizzazione di base, perché solo un patrimonio di conoscenze sufficientemente ampio e condiviso può costruire quel clima culturale che dà ancora più forza alle normative e soprattutto è in grado di influire sull'agire dei singoli. Occorre promuovere una cultura della cura¹² che si opponga alla cultura scarto¹³, alla convinzione che sacrificare una parte di umanità – magari riducendo i costi del lavoro¹⁴ – sia un prezzo accettabile sulla via del progresso tecnologico. È un passaggio necessario al fine di decolonizzare l'immaginario comune dalla narrazione salvifica che attribuisce alle macchine il potere di risolvere tutti i problemi che siamo chiamati ad affrontare e dal mito di una «*una crescita infinita o illimitata*»¹⁵.

Tra gli scenari apocalittici, che immaginano futuri distopici, e le concezioni salvifiche, che fanno della tecnologia la soluzione ad ogni problema, si apre la via mediana di un impiego delle IA che sia orientato a migliorare le condizioni dell'umanità, nella misura in cui esse siano soggette ad un'adeguata forma di controllo a livello culturale, sociale e politico.

Non si tratta di demonizzare le IA, ma di evidenziare come queste macchine dovranno rimanere centrate sull'uomo e sulla difesa dei suoi diritti, affinché si possa evitare che «*l'unicità della persona venga identificata e ridotta a un insieme di dati*»¹⁶.

¹² Cfr. FRANCESCO, *Lettera enciclica Laudato si' sulla cura della casa comune*, 231.

¹³ FRANCESCO, *Discorso al Corpo diplomatico accreditato presso la Santa Sede*, 13 gennaio 2014.

¹⁴ FRANCESCO, *Lettera enciclica Fratelli tutti*, 20.

¹⁵ FRANCESCO, *Lettera Enciclica Laudato Si'*, 106.

¹⁶ *Ibid.*, 5.
