Dichiarazione a tutela dell'Intelligenza Umana[1] (HI).
Statement Pro Human Intelligence (HI)

Le ragioni di questa Dichiarazione

L'avvento della rivoluzione digitale e della sua più emblematica innovazione, cioè la così detta Intelligenza Artificiale Generativa, esigono un ripensamento del perimetro concettuale di quella Intelligenza Umana (HI) che ci differenzia non solo da altri organismi viventi con i quali condividiamo una basilare Intelligenza Biologica (BI), ma anche dalle macchine basate sull' Intelligenza Artificiale (AI) . Una forma artificiale di ipotetica intelligenza che alcune correnti di pensiero propongono non solo come complementare a quella umana ma anche competitiva con essa e, in alcuni casi estremi, sostitutiva.

La letteratura scientifica sul tema dimostra ampiamente che l'HI e AI condividono solo alcune caratteristiche ma non tutte e la seconda assume la prima come modello strutturale, basti pensare alle reti neurali artificiali. Tuttavia, non abbiamo ancora conoscenze sufficienti dal punto di vista scientifico per spiegare in modo esauriente ed univoco nemmeno come funzioni il modello ispiratore umano, cioè cosa si intenda per "cervello", "mente" e "coscienza". Ci sono diverse teorie in proposito e la ricerca sperimentale è ancora in pieno sviluppo. Allo stato attuale questi concetti richiamano tutti un "paradigma della complessità" che viene affrontato da molteplici discipline incrociando scienze umane e scienze naturali.

Pertanto oggi si pone in termini urgenti il tema della **tracciabilità intellettuale** di prodotti e servizi realizzati con il contributo totale o parziale dell'AI fornendo agli utenti un'adeguata, esplicita e preventiva informazione, cioè mettendoli in grado di **discernere le differenze**. Così come esigiamo etichette per rendere trasparente origine e natura dei cibi che alimentano il nostro corpo, è altrettanto legittimo pretendere questa tracciabilità per prodotti e servizi che alimentano la nostra mente e le nostre esperienze esistenziali.

1 HI = Human Intelligence, in italiano Intelligenza Umana (IU).

Discernimento e obbligo di trasparenza.

Non a caso “discernimento” è un concetto, declinato in varie forme, che attraversa tutta l’Enciclica Magnifica Humanitas pubblicata da Papa Leone XIV lo scorso 25 maggio e dedicata al tema della custodia della persona umana nell’epoca dell’intelligenza artificiale:

https://www.vatican.va/content/leo-xiv/it/encyclicals/documents/20260515-magnifica-humanitas.html.

In particolare l’Enciclica, pur apprezzando i notevoli progressi della tecnologia, tra i quali spicca oggi l’IA, ed ai quali non si oppone pregiudizialmente, richiama tuttavia “*ad un discernimento morale e sociale che custodisca il primato della persona, affinché sia sempre l’intelligenza umana, con la sua coscienza e la sua libertà, a guidare le innovazioni tecniche e a stabilirne con responsabilità l’uso e i limiti.*”

Parole che risuonano con il laico Regolamento Europeo sull’Intelligenza Artificiale (EU AI ACT) del 2024 che all’Articolo 50 del Capo IV recita: “*I fornitori garantiscono che i sistemi di IA destinati a interagire direttamente con le persone fisiche sono progettati e sviluppati in modo tale che le persone fisiche interessate siano informate del fatto di stare interagendo con un sistema di IA, a meno che ciò non risulti evidente dal punto di vista di una persona fisica ragionevolmente informata, attenta e avveduta, tenendo conto delle circostanze e del contesto di utilizzo*.” Per quanto attenuato nella seconda parte dell’articolo, viene sancito un diritto alla trasparenza a garanzia dei cittadini. Un obbligo di trasparenza al quale alcuni grandi player vorrebbero assolvere mediante agenti software in grado di “discernere” le differenze attraverso un “watermark” invisibile all’occhio umano. Ma questa opzione sarebbe davvero una soluzione del problema o una sua ingegnosa elusione?

Il processo di delega

Nonostante il richiamo ad una sorveglianza critica nell’impiego dell’Intelligenza Artificiale dichiarato dal Regolamento Europeo, nella pratica sociale assistiamo ad un veloce, inarrestabile e progressivo processo di “delega” di decisioni o pre-decisioni - tradizionalmente appannaggio di istituzioni umane - a macchine basate sull’AI. Fenomeno che non tiene in debito conto gli errori sistematici compiuti dall’AI (allucinazioni), i pregiudizi cognitivi (bias) annidati nei data-set di addestramento e, in generale, il fatto che queste macchine agiscono senza comprendere (*agere sine intelligere*), elaborano cioè simboli secondo regole formalizzate (algoritmi) senza comprendere il significato di tali simboli, come nel caso dei chatbot conversazionali che usano il linguaggio. Agiscono quindi secondo principi di correlazione statistica che tendono a prevedere la probabilità di occorrenza di una stringa di simboli dopo un’altra stringa di simboli. E’ a tutt’oggi indimostrato che tale procedura sia equivalente o assomigli a quella messa in atto da una mente umana nell’elaborare sia le informazioni provenienti dall’interno (cervello e corpo) sia quelle provenienti dal mondo esterno (ambiente).

Inoltre questo processo di delega non è “politicamente neutrale” perché favorisce un progressivo scivolamento del governo delle nazioni da istituzioni nazionali e internazionali che ne

rappresentano i popoli ad oligarchie tecnocratiche “esterne” difficilmente controllabili guidate da interessi meramente economici e spesso divergenti rispetto al perseguimento di quel “bene comune” fortemente richiamato nell’Enciclica Magnifica Humanitas.

Fine di un habitat antropocentrico?

Inoltre, come affermato dai filosofi e antropologi più critici, questa rivoluzione digitale tende a realizzare anche una nuova architettura dell’habitat umano. In altre parole, spinge affinché siano gli esseri umani ad adattarsi ad ambienti fisici e culturali a misura di “macchine intelligenti” e non viceversa. Un’architettura che investe la vita quotidiana, dalla domotica alle smart city, dove a dettare le modalità di interazione tra essere umano e ambiente e tra gli stessi esseri umani sono gli “algoritmi” che fanno funzionare le “macchine intelligenti” e non gli “antroporitmi” che garantiscono il buon funzionamento (salute) e il benessere psicologico delle persone.

Fallacia del Test di Turing

Anche il così detto “Test di Turing” escogitato per verificare se una macchina interrogata da un input umano (prompt) sia in grado di produrre risposte (output) indistinguibili da quelle che nelle medesime circostanze fornirebbe un essere umano, è stato fortemente criticato poiché non proverebbe una capacità di “pensiero intelligente” bensì la capacità di produrre sequenze di simboli coesi e coerenti con un insieme di regole formali. L’equivalenza degli output non prova l’equivalenza dei processi di elaborazione né, tantomeno, l’esistenza di intenzioni e scopi coscienti. Così come una calcolatrice elettronica programmata per eseguire calcoli matematici non può che produrre risultati di operazioni matematiche indifferenti agli scopi del calcolo e alle entità del mondo naturale a cui tali calcoli vengono applicati.

Le macchine non sono organismi biologici e pertanto non possono condividere, per definizione, né la caratteristiche della vita biologica (tendenza all’omeostasi, alla riproduzione ecc.) né la specifica “condizione umana” che rappresenta l’orizzonte di senso da cui scaturisce il significato di ogni produzione simbolica.

Il pericolo di una posizione ineluttabilista.

L’avanzare di questi utilizzi ormai onnipervasivi dell’AI, sostenuta da pressioni economiche molto forti di potenti player internazionali, sembra diffondere nell’opinione pubblica ed anche tra alcuni decisori politici una postura culturale arrendevole e rinunciataria che definiamo **ineluttabilismo**. Si sta diffondendo cioè un pericoloso convincimento circa l’inevitabilità di questa evoluzione, soprattutto tra le giovani generazioni post-digitali, che provoca l’arretramento dello lo spirito critico su una posizione molto difensiva secondo la quale l’unico obiettivo realistico, a tutela del genere umano, può essere quello di mitigare con qualche correttivo normativo gli effetti negativi della “rivoluzione digitale” sulla vita delle comunità e delle persone. Posizione agli antipodi del vibrante appello di Papa Leone XIV: “Disarmare l’IA”.

Cui prodest? A chi giova?

Una domanda fondamentale che viene poco approfondita anche negli innumerevoli dibattiti intorno alla rivoluzione digitale e all’AI, è la seguente: chi ne trae vantaggio?

Quando ci si sofferma su questo aspetto, fondante rispetto a qualsiasi altra questione, i fautori di uno sviluppo tecnologico "senza vincoli ostacolanti" argomentano che l'umanità intera ne avrà molti vantaggi.

Tuttavia questi vantaggi, da cui originano enormi profitti economici monopolizzati da una ristretta aristocrazia tecnologica, non sono equamente distribuiti tra coloro che lavorano per rendere possibile l'Intelligenza Artificiale e nemmeno tra i potenziali fruitori; inoltre comportano un notevole impatto sull'eco-sistema in quanto l'AI consuma grandi quantità di risorse naturali (acqua ed energia). Pertanto, i "danni collaterali" da AI svantaggiano tutti.

Inoltre, quando li esaminiamo più da vicino, ci rendiamo conto che questi vantaggi vengono collocati più nel futuro che nel presente e sono sovente circoscritti a determinati ambiti dove c'è la necessità di elaborare grandi quantità di dati per agevolare decisioni umane, velocizzare la realizzazione di prodotti e migliore l'erogazione di servizi; ad esempio il trattamento di testi, la traduzione interlinguistica, la scrittura di codici di programmazione; la produzione di materiali audiovisivi; l'applicazione di modelli meteorologici, l'ottimizzazione di sistemi di agricoltura; la ricerca medica sia in ambito preventivo che terapeutico; il supporto alle persone con disabilità; le telecomunicazioni. Nonostante questa circoscrizione degli ambiti applicativi genuinamente vantaggiosi, la rivoluzione dell'AI investe già in modo profondo tutti gli aspetti della vita quotidiana di miliardi di persone creando anche bisogni indotti e imponendo nuovi stili di vita, a partire da una nuova appendice fisica dalla quale ci separiamo raramente, lo smartphone.

Questa scatola magica inghiotte persone di ogni età aprendo le porte ad un mondo virtuale e piacevole, stimolatore inesauribile di dopamina, dove trascorrere molta parte del proprio tempo di vita. Tuttavia, siccome il tempo di vita non è infinito, ciò che si aggiunge da una parte si sottrae dall'altra. Pertanto, il tempo passato nella dimensione virtuale è una sottrazione del tempo speso nella dimensione reale. Abbiamo prove che questa sottrazione aumenti il grado di benessere fisico e spirituale degli esseri umani?

"C'è un tempo per ogni cosa" si legge in un antico libro sapienziale, l'Ecclesiaste. Non stiamo correndo il rischio che alla fine ci sia il tempo "solo per una cosa"?

A ciò bisogna aggiungere che tale ipotetico progresso tecnologico non è affatto equidistribuito se è vero che circa un terzo della popolazione mondiale non ha accesso a internet e dispositivi digitali (Digital Divide)[2].

Ferita narcisistica o incommensurabilità dei due tipi di intelligenza?

Spesso i transumanisti obiettano ai tecno-scettici rispetto allo sviluppo dell'AI, che dopo le tre grandi ferite narcisistiche della storia, la rivoluzione copernicana (*la terra e l'uomo non più al centro dell'universo*) , la rivoluzione darwiniana (*l'uomo non più creatura prediletta ma semplice anello della catena evolutiva animale*) e quella freudiana (*la ragione non più dominatrice assoluta della vita psichica*), tale scetticismo nascerebbe da un perdurante pregiudizio cognitivo incarnato da una postura culturale che può essere definita primatismo dell'umano. In altre parole, sarebbe

2 Dati forniti da ITU (International Telecommunication Union), organismo delle Nazioni Unite. https://www.itu.int/en/mediacentre/Pages/PR-2024-11-27-facts-and-figures.aspx

soltanto questo pregiudizio ad impedire a molti di accettare una quarta ferita, cioè l'avvento di una super intelligenza postumana - evento ipotetico spesso definito "singolarità tecnologica" - che sorpasserebbe per potenza e capacità la HI o ne cambierebbe radicalmente i connotati ibridando esseri umani e dispositivi tecnologici (cyborg). Tuttavia, questa prospettiva ignora completamente l'incommensurabilità delle due intelligenze. Una "qualitativa" vincolata dagli "antroporitmi" di un essere senziente, cosciente e mortale, caratterizzato da universali tensioni spirituali, e una "quantitativa" basata sugli "algoritmi" ma dipendente da un indimostrato paradigma riduzionistico secondo il quale qualsiasi aspetto della vita umana può essere oggetto di un'euristica computazionale.

Creatività umana e artificiale

Non è la prima volta che l'umanità si confronta con la necessità di distinguere e riconoscere l'unicità degli esseri viventi. Già alla fine del secolo scorso, con la nascita di Dolly, il primo mammifero clonato da una cellula adulta, emerse una riflessione fondamentale sul rapporto tra originale e copia, tra identità biologica e replicazione tecnologica. Oggi quella stessa domanda si ripropone in una nuova dimensione: non più soltanto sul piano genetico, ma sul terreno della creatività, dell'invenzione e della produzione culturale.

Se la clonazione ha posto il problema del riconoscimento dell'individuo biologico, l'era dell'intelligenza artificiale pone anche il problema del riconoscimento dell'individuo creativo. Come possiamo identificare, tutelare e valorizzare il contributo originale di una persona all'interno di processi sempre più automatizzati? Come possiamo garantire che l'invenzione, l'opera dell'ingegno, l'intuizione artistica o scientifica mantengano un legame riconoscibile con il loro autore?
Questa iniziativa nasce anche da tale esigenza: costruire strumenti culturali, giuridici e tecnologici capaci di preservare la riconoscibilità dell'apporto degli esseri umani nei processi creativi e innovativi, affinché il progresso tecnologico continui a svilupparsi nel rispetto dell'identità, della dignità e della responsabilità degli autori e degli inventori.