

COMMISSIONE DI STUDIO
DELLE CHIESE BATTISTE, METODISTE E VALDESI IN ITALIA
PER I PROBLEMI ETICI POSTI DALLA SCIENZA

INTELLIGENZA ARTIFICIALE.
NOTE PER UNA RIFLESSIONE ETICO-TEOLOGICA
IN PROSPETTIVA PROTESTANTE.

INDICE

Introduzione

1. IA: sfide e opportunità.....	4
2. Etica dell'IA: nozioni di base.....	7
2.1 Esempi di criticità	
3. AI Governance e legislazione internazionale.....	12
4. Come l'IA interroga la Teologia.....	15
4.1 <i>Creatura in relazione</i>	
4.2 <i>Soggetto etico</i>	
4.3 <i>Responsabilità in atto</i>	
5. Conclusioni.....	19
6. Glossario.....	20

Introduzione¹

L'intelligenza artificiale è ormai parte della nostra esperienza quotidiana, utilizzata in modo consapevole, ad esempio attraverso l'uso degli assistenti virtuali dei nostri dispositivi portatili o domestici, nelle nostre interazioni con *chatbot* per accedere a servizi, cercare informazioni, ricevere assistenza durante un viaggio e innumerevoli altre attività, sia in modo inconsapevole quando diventiamo fonte di dati con cui le intelligenze artificiali vengono addestrate per svolgere svariati compiti e costruire prodotti il cui utilizzo non è evidente all'utente finale non esperto.

Le diverse generazioni hanno mediamente un rapporto differenziato rispetto a queste innovazioni tecnologiche di grande impatto sulla nostra vita: i più anziani osservano con preoccupazione gli sviluppi e la pervasività di questo nuovo modo di fare ogni cosa, mentre le generazioni più giovani se da un lato sembrano già adattate a un modello di vita in cui le Intelligenze Artificiali² sono un dato presente e normalizzato, dall'altro appaiono più vulnerabili ed esposte agli effetti indesiderati dell'utilizzo massiccio di questa tecnologia potente e talvolta ambigua.

La differente risposta delle generazioni costituisce infatti la conferma rispetto alla previsione della crescente difficoltà a tenere il passo dell'avanzamento tecnologico da parte degli esseri umani che si verranno a trovare sempre più fortemente in ritardo nella capacità di adattamento ai modelli di vita e lavoro resi possibili dal perfezionamento di queste macchine³.

Gli interrogativi etici sollevati da una tecnologia che non solo modifica il nostro modo di lavorare, di divertirci, di socializzare, di informarci, di muoverci, di lavorare, di curarci e di tessere relazioni, ma il nostro stesso modo di pensare noi stessi sono molteplici e costituiscono una sfida per il pensiero. Si ripropone una riflessione già emersa molti decenni fa, agli albori della Bioetica: è possibile trovare risposte etiche a problemi nuovi con i modelli ereditati dalle tradizioni più antiche? O piuttosto l'impatto della tecnologia sulla nostra vita implica anche una profonda revisione delle categorie essenziali del pensare e dell'agire?

Le questioni sollevate dalla presenza sempre più importante di IA nelle attività quotidiane sono molteplici e riguardano certamente gli aspetti etici dell'utilizzo di questa tecnologia, con

¹La Commissione ha condiviso all'unanimità l'orientamento e l'impostazione del presente testo.

Estensore del Documento: Ilenya Goss. Hanno partecipato al percorso di studio, ricerca, discussione e revisione Anita Ammenti, Libero Ciuffreda, Stefano Bertuzzi, Roberto Labianca, Massimo Marottoli, Silvia Rostain.

² Nel presente testo si è scelto di adottare la forma italiana IA, ma frequentemente è possibile trovare l'abbreviazione inglese AI anche all'interno di testi redatti in Italiano.

³ Anche senza approfondire il concetto di "singolarità tecnologica" secondo il quale vi sarebbe un punto di sviluppo a partire dal quale il controllo delle IA diventerebbe problematico, già oggi vediamo talvolta un rapporto caratterizzato da incomprensione, sospetto e paura, proprio mentre il potenziale di questa tecnologia si dispiega venendo a costituire un enorme accumulo di potere nelle mani dei pochi in grado di governarne i processi a fronte di una massa di utilizzatori inconsapevoli.

specificità legate a ciascun settore di applicazione. Il desiderio umano di costruire macchine capaci di svolgere compiti paragonabili a quelli per cui è chiamata in causa la nozione di *intelligenza*, non ultima l'interazione con la macchina in modalità sempre più simili a quelle interumane, costituisce un aspetto affascinante e inquietante su cui la riflessione filosofica e teologica ha iniziato a lavorare.

A questo proposito si intende offrire un breve contributo generale alla riflessione su questo tema vasto e in continua evoluzione, cercando di chiarire alcuni aspetti di ciò che oggi chiamiamo Intelligenza Artificiale e dei suoi utilizzi, facendo emergere le criticità più importanti e tracciando una via al pensiero teologico della “rivoluzione digitale” con tutte le sue rilevanti e non del tutto prevedibili conseguenze sulla vita degli esseri umani e del pianeta.

Si tratta di un testo inaugurale, destinato forse in alcune parti a una rapida obsolescenza, tuttavia necessario per dare avvio a un lavoro indispensabile nel quadro della responsabilità formativa e informativa che le Chiese accolgono come mandato.

1. IA: sfide e opportunità.

Si inizia a parlare di Intelligenza Artificiale negli anni Cinquanta del Novecento⁴, quando il potenziamento della capacità computazionale dei calcolatori, l'accesso a quantità crescenti di dati e la loro memorizzazione realizzano nuove applicazioni nell'area informatica. Ma la ricerca sull'Intelligenza Artificiale, cioè la produzione di macchine in grado di eseguire compiti che quando vengono svolti da esseri umani riteniamo necessitino di intelligenza⁵, riguarda anche altre scienze come la matematica e la statistica, la robotica, lo studio e le applicazioni delle reti neurali. Le competenze messe in gioco sono molteplici e diversi gli ambiti interessati.

La prima fase di ricerca intorno al “computer intelligente” vede l'articolarsi di teorie diverse sulle sue possibili evoluzioni: da un lato la teoria “forte” che ipotizzava la possibilità di macchine pensanti, dall'altro la tesi ben più realistica di grandi potenziali di sviluppo, ma al di fuori di ciò che chiamiamo in senso proprio “intelligenza”. In effetti nei decenni successivi la seconda prospettiva mostra il proprio realismo e segna una fase della ricerca in cui mentre alcuni orizzonti appaiono semplicemente fantascientifici si fa strada la consapevolezza delle vere possibilità di sviluppo più promettenti.

⁴ Si fa risalire la storia dell'IA al 1956 quando un progetto di ricerca di un gruppo di matematici americani presso il Dartmouth College usa per la prima volta questa espressione. Naturalmente si giunge a questo sviluppo grazie a un lavoro iniziato almeno venti anni prima.

⁵ Il test di Turing stabilisce che una macchina può esser detta capace di comportamento “intelligente” se le sue risposte sono indistinguibili dalle risposte di un essere umano alle stesse domande. La discussione sull'uso del termine “intelligenza” si protrae fino a oggi anche se ormai entrato nell'uso comune: uno degli effetti dell'uso per alcuni decisamente improprio e fuorviante della parola è una riconsiderazione anche filosofica del suo complesso significato.

Oggi la ricerca legata alle Intelligenze Artificiali Generative (AGI) consente per certi aspetti di lasciare aperta la possibilità di un'ulteriore evoluzione in direzione dell'originaria teoria forte, ma soprattutto ripropone la questione di cosa sia l'*intelligenza*, dal momento che oggi si concorda sul continuare a utilizzare questa espressione pur sapendo che in senso proprio la macchina non può essere definita intelligente⁶.

Alle IA in uso oggi si giunge attraverso un percorso che ha da un lato ridimensionato alcune aspettative iniziali, ma dall'altro ha moltiplicato esponenzialmente le possibilità di utilizzo in ambiti molto diversi. Menzioniamo tre grandi aree di applicazione che incontriamo con alta frequenza:

a) comunicazione, algoritmi, social media:

una parte importante della comunicazione, compresa quella politica con ricadute fondamentali sulla finanza, avviene attraverso i social media che sono produttori, raccoglitori e organizzatori di una mole crescente di dati gestita attraverso algoritmi in grado di modificare sensibilmente il quadro sociale e politico in cui viviamo. La possibilità delle IA di produrre immagini, ma anche contenuti di vario tipo, dai testi alla musica, alla riproduzione di specifici timbri vocali, apre la via a utilizzi che devono essere attentamente conosciuti e controllati;

b) chatbot e Large Language Models:

lo sviluppo dei Modelli Linguistici di grandi dimensioni (Large Language Models LLM) ci consente oggi di conversare con assistenti virtuali e di chiedere assistenza alle macchine per la preparazione di testi, la traduzione, la ricerca di informazioni organizzate: le possibilità di interazione non si limitano tuttavia al porre a disposizione dell'utente una enorme banca dati, ma nel mettere in atto vere e proprie conversazioni in cui la capacità delle IA di imitare il linguaggio diventa oggetto di osservazione e talvolta di preoccupazione per gli scenari di utilizzo che si prospettano⁷;

c) temi bioetici e Medicina.

l'area di interesse classico della Bioetica incontra il tema dell'IA in diversi aspetti, da quelli più innovativi legati alla cura dell'ambiente, a quelli tradizionali connessi al mondo della ricerca biomedica e della cura, fino alle questioni aperte riguardanti la sostenibilità energetica e ambientale dei supporti hardware di questa tecnologia.

⁶ Non sarà qui affrontata la tecnologia del computer quantistico (siamo infatti nell'era definita NISQ-*Noisy Intermediate Scale Quantum*- e all'ambito di ricerca QML -*Quantum Machine Learning*), perché si tratta di fase teorica e prototipale.

⁷ In ambito di ricerca avanzata troviamo anche gli LCM (*Large Concept Models*) e la loro integrazione con i World Models, che superano il livello di previsione statistica per approdare alla simulazione del funzionamento della realtà.

I diversi ambiti della Medicina sono ampiamente coinvolti nell'utilizzo delle IA a diversi livelli: nella clinica vediamo applicazioni che vanno dalla diagnostica per immagini, il primo campo in cui sono stati ottenuti risultati convincenti, all'oncologia, alla cardiologia, alla chirurgia, fino alle possibilità organizzative offerte dalla applicazione di IA a livello gestionale o nell'organizzazione dei percorsi personalizzati dei pazienti, anche attraverso la domotica medica.

Nell'ambito della ricerca biomedica vediamo le grandi potenzialità che vanno dalla progettazione di vaccini, alla selezione di molecole per nuovi farmaci, alla costruzione di modelli sperimentali.

Campo emergente e molto promettente, che tuttavia pone affascinanti interrogativi antropologici, è quello dell'area della biotecnologia applicata alle neuroscienze con lo sviluppo delle possibilità dell'interfaccia cervello-computer: si ripropongono sotto nuovi aspetti anche le questioni dello *Human Enhancement* e della neuroscienza computazionale⁸.

Le capacità sviluppate nelle IA riguardano la gestione dei dati non strutturati (consentendo profilazioni di vario utilizzo), le capacità percettive (sensori e programmazione per il riconoscimento di pattern specifici), le capacità di assistenza fisica tramite robot (assistente androide); l'insieme di queste capacità in diverse combinazioni può essere applicato a servizi diversi, che vanno dalle auto a guida autonoma, alla gestione di pazienti che necessitano di assistenza e monitoraggio a una gamma di possibilità ancora inesplorate e in crescita esponenziale.

Altri ambiti di utilizzo di IA sono lo studio dell'ambiente e la sua salvaguardia, il mondo dell'industria, quello delle assicurazioni e della finanza, del divertimento, fino a settori eticamente controversi come quello bellico o del controllo sociale. Ricordiamo anche l'impatto dell'IA nel campo dell'apprendimento, dell'assistenza psicologica o religiosa: imparare una lingua interagendo con un *coach* virtuale, confidarsi con un *chatbot* o rivolgersi a un *avatar* per conversare di fede è oggi alla portata di tutti.

A fronte di tanta varietà e duttilità di questa tecnologia è difficile sia prevedere temi e problemi che si proporranno anche tra poco tempo, sia raccogliere le questioni sotto un denominatore comune; tuttavia almeno un elemento generale che caratterizza le IA è quello di poter produrre risultati simili a quelli dell'azione, della decisione e del lavoro umano. In questo senso l'imitazione dell'intelligenza che oggi le macchine sono in grado di offrire consiste nel raggiungimento della stessa risposta, ma attraverso percorsi completamente diversi da quelli dell'intelligenza umana.

⁸ Per una informazione relativa allo Human Enhancement e una proposta di riflessione critica in chiave protestante sui temi legati al transumanesimo cfr. *Ragioni e limiti del potenziamento umano. Riflessioni sul ruolo sociale delle biotecnologie*, Doc.n.17/2014 Commissione bioetica BMV, <https://chiesavaltese.org/documenti/documento-studio-etica/>.

Quando parliamo di intelligenza umana comprendiamo nella parola non solo la capacità computazionale, ma anche caratteristiche nelle quali ci identifichiamo e che generano il fenomeno complesso della coscienza. *Instaurare relazioni o compiere azioni eticamente valutabili è proprio frutto della coscienza come centro focale dell'intelligenza umana.*

Se dunque è possibile individuare una questione etica generale legata alle IA essa si troverà non tanto negli specifici problemi che essa propone in ogni particolare campo di applicazione, ma nel fatto stesso di interagire con una macchina che imita e potenzia alcuni aspetti del lavoro dell'intelligenza di chi l'ha costruita: questo ci conduce verso scenari inediti tanto affascinanti quanto problematici.

2. Etica dell'IA: nozioni di base.

Negli ultimi anni la riflessione sulle questioni etiche legate all'utilizzo delle IA è diventata pervasiva: se ne occupano istituzioni di ricerca, di formazione, istituzioni religiose, il sistema legale, le aziende, le associazioni di consumatori: si tratta di un vero e proprio campo aperto in cui ci si imbatte a diversi livelli, accademici o operativi.

I primi problemi emersi rispetto all'utilizzo di questa tecnologia riguardano la *privacy* dei dati (la questione dei Big Data e della loro gestione e proprietà) e *l'equità* nella loro gestione, grazie all'emergere di studi sui *bias*⁹ di apprendimento delle macchine; a questi primi temi si sono aggiunte le riflessioni intorno alla *trasparenza* del funzionamento delle IA, alla loro *sicurezza*, al *controllo* da parte umana e all'aderenza a *valori* condivisi nella loro progettazione e addestramento.

Così come le professioni sviluppano codici deontologici che stabiliscono criteri e principi significativi in uno speciale ambito applicativo, ma che sono inseriti in un quadro generale di etica condivisa, anche le applicazioni delle IA in diversi settori fanno emergere con maggiore o minore evidenza la necessità di applicare determinati principi perché l'utilizzo delle IA sia eticamente inquadrato in un modello di linee guida generali.

Senza poterci addentrare qui in ciascuno dei campi speciali di applicazione dell'IA è possibile descrivere orientamenti generali che indicano i limiti entro i quali l'utilizzo dell'IA può avvenire nel rispetto di un modello di vita umana improntato ad alcuni principi fondamentali: questa considerazione suggerisce le dimensioni globali del tema, che in effetti riguarda l'essere umano in

⁹ Cfr. Glossario. Il problema dei *bias* di addestramento è stato sollevato rispetto alle applicazioni di IA in campi come la diagnostica, il riconoscimento facciale, le assicurazioni ecc.: avendo addestrato la macchina con dati provenienti da una popolazione selezionata (ad esempio prevalentemente maschi bianchi di una specifica fascia sociale, economica, ecc...) le risposte ottenute alla presentazione di casi differenti erano scorrette riflettendo un pregiudizio. Questo ha evidenziato la questione etica dei potenziali effetti discriminatori dell'addestramento (*Deep Learning*). Ci sono *bias* introdotti durante la progettazione dell'algoritmo e *bias* intrinseci nella scelta dei dati.

quanto tale e non soltanto una categoria di professioni, dato che l'impatto di questa trasformazione coinvolge molteplici aspetti di ciò che costituisce oggi la nostra esperienza di "mondo".

Già da alcuni anni gli esperti¹⁰ hanno individuato alcuni principi etici generali particolarmente significativi per l'utilizzo dell'IA, condivisi dagli studiosi che si sono occupati del tema, che possono essere raccolti in poche parole-chiave:

- trasparenza,
- equità e non-discriminazione,
- robustezza tecnica e sicurezza,
- responsabilità,
- controllo e sorveglianza umana,
- benessere sociale e ambientale
- privacy e governance dei dati

Tali principi intercettano l'esigenza di evitare che i pochi soggetti, pubblici o privati, in grado di sostenere finanziariamente lo sviluppo di IA, lo facciano sotto la spinta esclusiva di fattori economici o di potere e controllo sulle società i cui effetti sono solo in parte prevedibili.

La necessità di focalizzare l'attenzione sulle criticità etiche è percepita a livello globale, come testimoniano le iniziative degli ultimi anni, anche in Europa dove nel 2018 nasce il primo *framework* etico, *AI4 People*, mentre a livello normativo ricordiamo la Raccomandazione UNESCO del 2021 e l'AI Act europeo del 2024¹¹.

Il lavoro che oggi pare urgente comprende alcuni compiti essenziali come:

- individuare criteri etici standardizzabili e traducibili a livello ingegneristico;
- istituzione di comitati o *ethics-by-design* per la valutazione etica *ex ante* dei sistemi IA ad alto rischio,
- impegno di formazione interdisciplinare per ingegneri, giuristi e filosofi, per poter promuovere una *AI Governance* condivisa.

¹⁰ AI-High-Level-Expert-Group (AI-HLEG), gruppo costituito dalla Commissione Europea nel 2018, formato da 52 esperti in contatto con diversi stakeholder, finalizzato a proporre linee guida etiche e legali per l'IA in Europa.

¹¹ UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 25 novembre 2021. Parlamento Europeo, *Artificial Intelligence Act*, 13 marzo 2024.

Si apre inoltre lo spazio e la necessità di una considerazione filosofica della natura dell'azione algoritmica: necessitiamo di studiare o trovare nuove categorie per pensare l'*intenzionalità*, l'*autonomia* e la *normatività* in ambito IA¹².

Così come nel *design* degli strumenti che utilizziamo rientrano criteri di sicurezza, costi, sostenibilità, così per quanto riguarda gli algoritmi entra in gioco la considerazione degli effetti e delle risposte che il sistema potrà fornire quando gli chiederemo di risolvere per noi un problema o di rispondere a una domanda. Una grande questione tecnica ed etica è infatti costituita dai *bias* di addestramento di cui si è già fatta menzione, ma anche di aspetti meno controllabili delle risposte delle macchine¹³. Si sta prendendo coscienza in questi anni anche di una prevalenza di figure professionali maschili tra coloro che studiano, sviluppano, addestrano le IA, e di una conseguente modellizzazione non neutra rispetto al genere che sarà necessario al più presto approfondire e documentare onde apportare le giuste correzioni¹⁴.

Il lavoro intellettuale è oggi fortemente segnato dall'uso di IA: nella giustizia, nell'insegnamento, nella gestione di molti aspetti dei processi diagnostici, terapeutici, nella ricerca scientifica, nel mondo assicurativo, nell'economia e nella comunicazione. L'interlocuzione con *chatbot*, l'utilizzo di traduttori, dell'interfaccia IA dei database consultabili in rete, degli assistenti virtuali, si trovano su una linea di confine in cui l'utilizzatore dello strumento e l'IA entrano in un rapporto che implica un processo di interazione costante (l'addestramento di un *chatbot* ad esempio) che imita l'interlocuzione interumana.

Il problema presenta due aspetti: da un lato l'interazione con la macchina "come se" fosse un interlocutore umano ha sicuramente un effetto di ritorno sull'umano; ma d'altra parte, e di rilevanza etica, resta aperta la questione del soggetto che controlla gli effetti di questo utilizzo. Via via che la decisione, la gestione dei dati e dei sistemi (dal semplice utilizzo di un termostato in grado di rilevare la nostra presenza in casa, a sistemi di controllo di grandi impianti di diversa natura) vengono delegati

¹² A questo proposito sarà utile ricordare che alcuni anni or sono in Italia la proposta di un termine per indicare l'interesse etico nel lavoro sull'Intelligenza Artificiale ha acquisito popolarità senza tuttavia riscuotere un vero consenso: la parola "algoritmica" starebbe a indicare il campo tecnico di studio finalizzato a includere l'etica nel processo di *design* degli algoritmi. Acquisito che per quanto potenti le macchine sono strumenti privi di un assetto valoriale autonomo, il soggetto etico rimane l'essere umano, il solo responsabile di tutto il percorso compiuto dall'IA, dal suo *design*, all'addestramento, alle richieste fino anche alle risposte che la macchina è stata messa in condizioni di fornire. In tal senso l'algoritmica risulterebbe essere semplicemente un modo sintetico per indicare che un criterio etico entra in gioco insieme alle competenze tecnico-informatiche quando si programma un algoritmo. Il termine sarebbe comunque travisato se inteso come affermazione di una "eticità" intrinseca degli algoritmi, poiché la soggettività etica non è spostabile al di fuori dell'umano, l'unico a cui sia attribuita anche una eventuale imputabilità giuridica.

¹³ Pur non essendoci consenso sul tema va ricordata la questione della *black box*, quell'aspetto di non-trasparenza e non completa prevedibilità delle risposte dell'IA.

¹⁴ La questione dei *bias* di addestramento è complessa: a titolo di esempio anche nel training di IA destinate a fornire assistenza in ambito spirituale/religioso la ridotta capacità delle minoranze di fornire i propri testi, dati, ordinamenti, informazioni finisce con il condizionare la risposta del modello determinando la sottorappresentazione di determinati contenuti e gruppi.

alla macchina gli effetti di lungo periodo si allontanano sempre più dalla possibilità di un reale controllo, generando una dipendenza della quale non si comprende al momento la portata.

2.1. Aree generali di criticità.

L'etica applicata all'IA ha come obiettivo la progettazione, la produzione e l'utilizzo di sistemi che siano coerenti con i diritti fondamentali dell'uomo; dal momento che i sistemi automatizzati hanno un forte impatto sociale, in Europa vi è un forte impegno perché questa potente tecnologia sia gestita eticamente, sia a servizio del bene comune e non appannaggio di privilegiati o peggio utilizzata minando giustizia e libertà e mettendo a rischio i sistemi democratici. Sono state individuate alcune aree generali sensibili in cui occorre un controllo mirato del lavoro dell'IA:

a) bias e discriminazione. Come si è visto (cfr. note 9 e 14) la progettazione dell'algoritmo e l'addestramento della macchina possono incorporare distorsioni dovute a incompletezza oppure a selezione a monte dei dati. Questo può generare risposte errate o comunque discriminatorie, evidenziando un problema etico di giustizia e di equità.

b) privacy e sicurezza. La gestione dei *Big Data* è diventata molto presto oggetto di preoccupazione etica: il tema è complesso e comprende la profilazione inconsapevole, la vendita di dati, ma anche rischi di utilizzo illecito da parte di chi li raccoglie, li detiene e li gestisce. Si aprono a questo proposito molte sotto-aree di interesse etico, ad esempio quelle legate alla possibilità di cancellazione della traccia digitale che lasciamo.

c) responsabilità, trasparenza e controllo. L'utilizzo di IA nei campi più disparati del lavoro o del tempo libero comporta la possibilità di errore da parte della macchina e il conseguente tema della responsabilità civile o eventualmente penale per il danno arrecato. Anche il tema della trasparenza e della controllabilità delle risposte offerte dall'IA, la possibilità di risalire a un responsabile, sono rilevanti per un uso etico della macchina e la gestione della delega.

d) lavoro e automazione. In quest'area si apre il tema ampio dei cambiamenti che iniziamo a osservare in diversi settori lavorativi in cui la macchina gestita da IA potrà sostituire il lavoro umano. Questo solleva complessi problemi etici e di gestione sociale, di distribuzione della ricchezza e conseguentemente di organizzazione politica.

e) lavoro umano. Per addestrare le IA possono venire utilizzati lavoratori, spesso sottopagati e nei paesi in via di sviluppo, sottoposti in alcuni casi a ritmi di lavoro usuranti ed esposti a immagini e altri contenuti fortemente disturbanti, con potenziali effetti psicologici e fisici negativi.

f) disuguaglianza economica e divario digitale. Le possibilità offerte dalle IA comportano la capacità di sostenere dei costi che naturalmente possono generare disuguaglianza sociale ed esclusione di fasce di popolazione più fragile sia dal punto di vista economico sia dal punto di vista dell'istruzione e della formazione.

g) proprietà intellettuale e produzione di contenuti artistici. La capacità dell'IA di raccogliere dalla rete dati e informazioni, immagini, audio, e di avere accesso a vasti database fa sì che non solo la possibilità di plagio aumenti, ma che la protezione del lavoro intellettuale e la proprietà sia messa a rischio.

h) impatto educativo. L'utilizzo di IA fin dall'età scolare sta già mostrando modifiche non solo nei comportamenti dei bambini e dei ragazzi, ma anche potenziali effetti neurobiologici di cui non conosciamo ancora la portata. Gli effetti di natura psicologica sono già evidenti in alcuni comportamenti diffusi legati a utilizzo costante di *device* e di IA.

i) sostenibilità ambientale. Spesso si trascura di pensare che l'IA necessita di strutture fisiche per funzionare e che per quanto appaia immateriale ha invece un versante di consumo non solo energetico, ma anche di spazi fisici e di risorse naturali. Questo solleva una questione importante di sostenibilità ambientale, tema della Bioetica nuova che si va costruendo in questi anni.

Non si tratta di un elenco esauriente, ma di aree problematiche in cui sono già stati osservati risultati eticamente discutibili o francamente non accettabili. A questa preoccupazione si aggiungono le possibilità di utilizzo intenzionalmente anti-etico legate ai rischi per la libertà, e gli utilizzi militari¹⁵.

La discussione etica intorno allo sviluppo dell'IA, della robotica e di tutte le applicazioni che emergeranno deve tener conto del fatto che l'orizzonte disegnato da questa rivoluzione è molto più ampio di quanto possa apparire: in Occidente la concentrazione del potenziale economico e tecnologico di sviluppo in pochi soggetti privati e la riproposizione di forme nuove di colonialismo¹⁶, mentre in Oriente una capacità di sviluppo che può contare sull'organizzazione compatta dello Stato (Cina), fanno sì che le prospettive superino il piano della semplice regolamentazione mettendo in gioco forze difficili da governare. In tale quadro geopolitico l'Europa potrebbe forse avere un ruolo importante per uno sviluppo etico dell'IA.

¹⁵ Il 18 aprile 2026 è stato pubblicato da Alex Karp il Manifesto di Palantir: in esso emerge una prospettiva globale sullo sviluppo dell'IA riassumibile in un superamento dell'assetto geopolitico attuale e in un radicale cambiamento di prospettiva che sostituisce il potere tecnologico alle regole del diritto e della democrazia. Al centro del progetto c'è l'uso militare dell'IA e l'impiego per il controllo della popolazione in vista di un modello di Occidente in grado di contrapporsi alla potenza tecnologica, economica e militare della Cina.

¹⁶ Il riferimento è al problema di cui si fa menzione al punto e) a pag. 10.

3. AI Governance e legislazione internazionale¹⁷.

La disciplina giuridica dell'IA è in costante divenire ed è caratterizzata da un quadro eterogeneo nel quale coesistono fonti vincolanti, strumenti di *soft law* e iniziative pubbliche e private di coordinamento europeo ed internazionale.

Con il termine *AI Governance* si definisce il complesso sistema di norme, istituzioni e meccanismi di monitoraggio preposti a regolare l'intero ciclo dell'IA, dalla progettazione all'impiego. Tale architettura giuridica è intrinsecamente legata a un nucleo di valori fondamentali che attingono alla bioetica, alla tutela internazionale dei diritti umani, all'etica delle tecnologie e alla disciplina sulla protezione dei dati personali.

La necessità di standard comuni è stata avvertita inizialmente nel settore privato, al fine di assicurare gli operatori in ordine alla sicurezza nel suo utilizzo: oggi, lo sviluppo di sistemi di IA deve integrare i requisiti normativi e le esigenze di controllo sin dalla fase di progettazione.

Il primo aspetto problematico per la tutela universale dei diritti fondamentali della persona, in un contesto tecnologico privo di confini come quello della IA, risiede nell'assenza di una definizione di diritto fondamentale che sia omogenea, condivisa e sovrapponibile a livello internazionale.

In base agli standard vigenti, un'efficace *governance* dell'IA deve prevedere soluzioni operative in grado di presidiare l'intero ciclo di vita del sistema: dalla definizione del caso d'uso e dei centri di responsabilità, all'analisi del contesto applicativo, fino alla classificazione del rischio, garantendo un monitoraggio costante e una rigorosa supervisione umana (*human oversight*¹⁸ o *human-in-the-loop*¹⁹).

La Legge italiana che al momento norma l'uso dell'IA è la n. 132/2025²⁰ che introduce una disciplina nazionale distinta, ma complementare e coerente con il quadro europeo, riguardante la sperimentazione, l'adozione di sistemi algoritmici, lo sviluppo e l'applicazione di modelli IA nella Pubblica Amministrazione. È una legge che pone un preciso rapporto tra innovazione tecnologica, responsabilità umana e decisione amministrativa. La Legge si pone nel quadro del Regolamento UE 2024/1689 del Parlamento Europeo (*AI Act*).

¹⁷ *AI Act*. Principi, regole e applicazioni pratiche del Reg UE 1689/2024.

¹⁸ <https://artificialintelligenceact.eu/article/14>.

¹⁹ *Human-in-the-Loop* (HITL) si riferisce a un sistema o processo dove un essere umano partecipa alle operazioni, alla supervisione o al processo decisionale di un sistema automatizzato. Nel contesto dell'AI, HITL significa che gli esseri umani sono coinvolti a un certo punto del processo per garantire precisione, sicurezza, responsabilità. Ad esempio una terapia suggerita da un sistema diagnostico IA che viene erogata solo dopo la firma digitale del medico.

²⁰ Legge 132/2025, “*Disposizioni e deleghe al Governo in materia di intelligenza artificiale*”.

Il testo insiste sul principio antropocentrico come guida per l'uso dell'IA e sottolinea i principi di responsabilità, autonomia e trasparenza oltre a introdurre obblighi di tracciabilità e riconoscibilità dei modelli adottati. La Legge ha come obiettivo quello di rendere controllabili in sede giurisdizionale e valutabili in sede amministrativa tutte le decisioni che dovrebbero essere sempre riconducibili a un responsabile nel tentativo di evitare il cosiddetto “vuoto di responsabilità” causato dalla “opacità” degli algoritmi. Il principio cardine ivi espresso risiede ancora una volta nella centralità dell'essere umano: lo sviluppo e l'applicazione dei sistemi di IA devono sempre rispettare l'autonomia e il potere decisionale dell'uomo, assicurando la sorveglianza e l'intervento umano²¹.

Importante riferimento per lo sviluppo e la regolamentazione dell'IA in Italia è il documento *Strategia italiana per l'Intelligenza Artificiale*, un testo dinamico, approvato con cadenza biennale, per delineare gli obiettivi a lungo termine, gli investimenti, i principi e le azioni concrete per governare la transizione tecnologica nel rispetto dei valori nazionali ed europei²². Esso è caratterizzato da tre obiettivi principali: supportare l'innovazione e dare sostegno allo sviluppo imprenditoriale, promuovere la ricerca applicata e migliorare le competenze digitali, sviluppare una regolamentazione etica.

La strategia europea in materia di IA si basa su due pilastri normativi principali: il *Regolamento sull'Intelligenza Artificiale* (noto come *AI Act*) e il *Regolamento Generale sulla Protezione dei Dati (GDPR)*, supportati da una solida base di principi etici e da una crescente giurisprudenza della Corte di Giustizia dell'Unione Europea (CGUE).

L'*Artificial Intelligence Act*, approvato nel 2024²³, costituisce il primo quadro giuridico organico e vincolante in materia di IA: presenta un approccio basato su quattro categorie di rischio, inaccettabile, alto, limitato e minimo, e fissa obblighi proporzionati per garantire un elevato livello di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla *Carta dei diritti fondamentali* dell'Unione europea.

Il testo disciplina l'immissione sul mercato, la messa in servizio e l'uso dei sistemi di IA nell'UE, stabilisce divieti per usi contrari ai valori dell'UE (ad esempio uso di sistemi che manipolano il comportamento umano o sfruttano le vulnerabilità di gruppi specifici), stabilisce requisiti specifici, obblighi di trasparenza come l'obbligo di informare le persone che stanno interagendo con una macchina o con sistemi che generano contenuti (*deepfake*). Sono inoltre presenti misure a sostegno dell'innovazione, con un'attenzione particolare alle piccole e medie imprese e alle *start-up*, anche

²¹ Legge n. 132 / Capo I, Art. 3.

²² <https://www.agid.gov.it/it/notizie/pubblicato-il-documento-completo-della-strategia-italiana-per-lintelligenza-artificiale-2024-2026>.

²³ https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=OJ:L_202401689 . Reg. UE 2024/1689.

attraverso l'istituzione di spazi di sperimentazione normativa (*regulatory sandboxes*) per testare sistemi innovativi in un ambiente controllato²⁴.

L'AI Act si coordina con la *Carta dei diritti fondamentali dell'Unione europea* e con il *Regolamento Generale sulla Protezione dei Dati* (UE) 2016/679 (GDPR), il primo come base fondamentale per i principi etici e i diritti, il secondo come riferimento imprescindibile rispetto a una tecnologia che tratta e processa volumi ingenti di dati personali, garantisce trasparenza²⁵ e disciplina i trattamenti automatizzati e la profilazione²⁶ applicando principi di "responsabilizzazione" (*accountability*) e di sicurezza²⁷. L'ordinamento europeo si completa attraverso una pluralità di atti complementari (quali normative in materia di sicurezza dei prodotti, responsabilità civile e servizi digitali), che concorrono a delineare un ecosistema integrato di regole.

Negli Stati Uniti d'America troviamo invece un approccio più pragmatico che interviene con strumenti mirati, come gli ordini esecutivi, per affrontare problemi specifici, in particolare quelli emersi nel contesto dei trasferimenti internazionali di dati e delle esigenze di sicurezza nazionale.

Il modello cinese declina la protezione dei diritti in chiave di stabilità sociale e sicurezza nazionale e si pone chiaramente l'obiettivo di essere tra i leader mondiali nello sviluppo di questa tecnologia. Alla fine del 2025, la *Cyberspace Administration of China* (CAC) ha predisposto uno schema regolatorio volto a disciplinare i servizi di IA che simulano interazioni umane. Il documento, sottoposto a consultazione pubblica fino al gennaio 2026, si rivolge specificamente ai fornitori di sistemi che intrattengono dialoghi e utilizzano espressioni "emotive" assimilabili a quelle umane. Queste misure segnano un'evoluzione significativa nel paradigma giuridico cinese, poiché fondono in un unico quadro istanze eterogenee: dalla tutela dei minori e dei dati personali alla sicurezza nazionale e al controllo dei flussi informativi²⁸.

La riflessione che segue a una semplice revisione della situazione legislativa può solo sottolineare la difficoltà di governare una trasformazione che è per sua natura globale in un contesto di legislazioni differenti.

²⁴ https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=OJ:L_202401689 .Cpt_ VI artt. 57 e ss.

²⁵ Gli articoli 13, 14 e 15 GDPR obbligano il titolare del trattamento a fornire all'interessato informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento.

²⁶ <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/6264597> art. 22 cristallizza una garanzia fondamentale, sancendo il diritto dell'interessato a non essere assoggettato a decisioni basate esclusivamente su processi automatizzati che producano effetti giuridici o incidano in modo significativo sulla sua sfera personale;

²⁷ Cfr artt. 24 e 32 GDPR che impongono al Titolare del trattamento di adottare misure tecniche e organizzative adeguate al rischio per garantire e dimostrare la conformità al Regolamento.

²⁸ <https://www.altalex.com/documents/news/2026/01/29/diritto-ia-cina-nuova-frontiera-regolazione-digitale>.

4. Come l'IA interroga la Teologia.

Lo sviluppo dell'IA e le sue applicazioni costituiscono un'area di interesse per la riflessione filosofica non esclusivamente in chiave etica, ma anche per le rilevanti questioni legate all'autocoscienza degli esseri umani, alle domande sui diversi riduzionismi applicati alla nozione di intelligenza e più in generale all'identità degli esseri umani in quanto tali.

Se i problemi etici generali e applicativi di questa tecnologia possono essere affrontati con gli strumenti della filosofia morale, gli interrogativi che essa pone sul rapporto dell'essere umano con sé stesso, con gli altri e con Dio, e sul suo ruolo nel mondo hanno certamente un interesse di tipo teologico: le prime evidenze di un effetto profondo dell'IA sulla percezione che gli esseri umani hanno della realtà, i cambiamenti cognitivi e di stile relazionale dei più giovani, le prospettive di *Human Enhancement*, le ricadute sul vivere associato, politico, ma anche ecclesiale, impongono una riflessione accurata che sarà tra i compiti della Teologia del prossimo futuro.

Non si tratta di lasciarsi spaventare da questi effetti: ogni innovazione tecnologica ha portato mutamenti anche profondi nelle modalità di vita e nel pensiero umano, anche se certamente la portata del cambiamento prodotto dalle IA è tanto rivoluzionaria da toccare punti sensibili dell'autopercezione e della consapevolezza umana della propria identità e specificità.

4.1 *Creatura in relazione.*

La prospettiva in cui l'etica teologica protestante legge la ricerca scientifica e l'innovazione tecnologica è disegnata dal confronto costante con la Scrittura come luogo in cui lasciarsi interrogare dalla Parola di Dio e in cui trovare elementi che orientino le decisioni a tutti i livelli dell'esperienza.

L'*adam* biblico compreso come creatura²⁹ a cui è affidato il compito speciale di servire e coltivare la terra³⁰ facendone un "mondo" da abitare è l'essere la cui identità è legata fin dalla sua origine all'*imago Dei*³¹; questo dato che inaugura anche una visione antropologica orienta la riflessione intorno al rapporto con le IA, frutto del lavoro umano che ha un effetto retroattivo sull'umano. Non si tratta di un manufatto tra altri, ma di sistemi e interfacce che cambiano profondamente il rapporto con l'esperienza del mondo, di macchine che imitano l'agire degli esseri umani quasi come un potenziamento della nostra capacità trasformativa del mondo. A partire da

²⁹ Il riconoscimento della creaturelità dell'essere umano fonda l'orizzonte interpretativo biblico sia dell'identità sia della vocazione dell'*adam*: rispecchiandosi nel venire al mondo chiamati da una volontà di vita e da una sovrabbondanza di benedizione, e non dal caso o da una inesplicabile causalità, determina un tipo di esperienza e di interpretazione dell'esistenza improntato alla gratitudine e alla responsabilità.

³⁰ "Dio il Signore prese dunque l'uomo e lo pose nel giardino di Eden perché lo lavorasse e lo custodisse" (Gn 2:15, Nuova Riveduta).

³¹ "Dio creò l'uomo a sua immagine: lo creò a immagine di Dio; li creò maschio e femmina" (Gn 1:27, NR).

questa rifocalizzazione sulla relazione come costitutiva dell'identità umana osserviamo in prospettiva teologica le sollecitazioni del nuovo ambiente disegnato dalle IA.

L'identità dell'essere umano creato a immagine di Dio si definisce attraverso il rapporto con il Creatore dove il *colloquio* assume un ruolo centrale; destinatario della Parola che gli viene rivolta l'essere umano è reso capace di risposta:

-l'imitazione del colloquio che avviene con l'IA non va confusa con l'**incontro** che un dialogo interumano e un dialogo divino-umano (la preghiera, la liturgia) costituiscono: per quanto simile a una interazione reale quella con l'IA è un utilizzo che imita una relazione. Il fatto che accada oggi sempre più frequentemente che il bisogno umano di relazione (percepita tuttavia come complicata, talvolta frustrante e faticosa) cerchi risposta nell'uso della tecnologia come suo sostitutivo, dovrebbe richiamare l'attenzione sulla *vocazione umana all'incontro con l'altro*;

-l'essere umano posto fin dalle origini in relazione con il proprio Creatore che instaura la relazione dialogica trova vita nella **verità** e morte nella menzogna: la vigilanza sugli utilizzi dell'IA che rischiano di compromettere la formazione e l'informazione, ma anche il dialogo cercato come compagnia sostitutiva, si presenta come compito rilevante. La teologia ha gli strumenti per offrire l'approfondimento di ciò che significa vivere il rispetto della Parola e delle parole e proteggere dal degradarsi del colloquio a pura funzione incapace di istituire *relazione con la verità*³²;

-l'IA dal punto di vista della gestione dei dati ha una potenza non commisurabile alla capacità umana, tuttavia l'intelligenza umana può essere pensata soltanto in connessione con altri aspetti che vanno a costituire l'esperienza e la conoscenza: le emozioni e i sentimenti, le percezioni corporee, la creatività sono alcuni elementi che a insieme alle prestazioni dell'intelligenza "computazionale" concorrono a costituire il *fenomeno della coscienza*. Esplorare la differenza tra essere *qualcuno* e essere *qualcosa*, tra oggetti e soggetto è una sfida teoretica per il pensiero filosofico e teologico del presente.

-l'identità umana è legata al senso del limite, alla ricchezza di significato di ogni aspetto della vita, compresi quelli della fragilità, della perdita, e della capacità di maturazione di atteggiamenti di compassione, di cura, di valorizzazione: uno sviluppo dell'IA animato esclusivamente dall'intenzione di rimuovere il limite, di trasformare l'essere umano potenziandolo tecnologicamente solleva non solo la questione della giustizia (chi e come ha accesso a questo potenziale), ma comporta anche un

³² La relazione propria del discepolato cristiano si struttura intorno al parametro della permanenza nella Parola e alla libertà che ne consegue (Gv 8:31-32).

rischio di dis-umanizzazione, di perdita di elementi fondamentali che situano l'essere umano nella sua relazionalità creaturale.

4.2 Soggetto etico.

Nella prospettiva dell'etica teologica riformata l'essere umano è chiamato a essere interlocutore di Dio e responsabile del proprio agire nel mondo: a un primo livello di indagine l'uso dell'IA rientra nel paradigma dell'utilizzo di uno strumento utile a fini accettabili oppure eticamente riprovevoli. Tuttavia, a differenza di quanto accade con prodotti della tecnica meno complessi, con l'IA abbiamo di fronte macchine a cui ci si affida per prendere decisioni, con il rischio della delega del compito umano della scelta. Inoltre nella programmazione e nell'addestramento delle IA generative possono esservi *bias* che, come abbiamo visto, possono riflettersi in un funzionamento viziato dal punto di vista etico sotto vari aspetti: dalla giustizia, alla riservatezza dei dati, all'equità di trattamento, alla trasparenza.

La chiarezza sulla responsabilità del soggetto etico è un contributo importante alla discussione odierna e consente di mantenere alta l'attenzione sul tipo di IA che entrano in uso, sottolineando l'importanza etica e il significato teologico del controllo umano perché esse siano al servizio del massimo bene in un quadro di principi condivisi.

Mantenere viva la consapevolezza che dietro gli algoritmi e le applicazioni IA c'è il lavoro di esseri umani e contribuire a sviluppare maggiormente il senso etico della loro produzione e del loro uso si presenta come un compito ineludibile: in questo momento storico lo sguardo teologico sul fenomeno trasformativo dell'IA porta l'attenzione sulla necessità di custodire il senso dell'identità umana e di accogliere il mandato di non subire passivamente la trasformazione in atto, ma di essere parte attiva di un progetto in cui offrire un contributo decisivo.

Lo sguardo teologico consente inoltre di mantenere chiara la centralità della relazione come struttura portante dell'essere umano: la costruzione di relazioni interumane (fino a quelle sociali, politiche) corrispondenti alla volontà di bene, di vita, di salute, di pienezza e realizzazione, la gestione del conflitto etico attraverso la nozione di responsabilità per altri, diventano orientamento rispetto alle grandi questioni sollevate dall'utilizzo dell'IA.

Anche senza fare riferimento ad ambiti di uso problematici di per sé, come l'uso delle possibilità di riconoscimento per il controllo delle popolazioni, i rischi per le democrazie che sono già evidenti, oppure l'applicazione militare delle IA, una generale considerazione del funzionamento delle IA conduce a riconoscerne il potenziale altamente competitivo e i conseguenti rischi per la giustizia e per la libertà degli esseri umani e per la forma del vivere associato e le relazioni tra Nazioni.

4.3 Responsabilità in atto.

Le chiese vivono immerse nella realtà condivisa da tutte e tutti, per questo si pone oggi l'esigenza anche teologica di non essere semplici utenti inconsapevoli della nuova tecnologia, ma di essere portatori di un pensiero critico, testimoni di istanze etiche e soprattutto sviluppatori di utilizzi innovativi dell'IA.

Investire nella formazione perché sia possibile essere presenti attivamente nel processo in corso diventa oggi una priorità di cura verso gli esseri umani che assume anche valore politico: il mandato specifico di servizio in cui le chiese riconoscono la propria *mission*, e il Protestantismo con la peculiarità culturale che lo ha contraddistinto fin dalle origini, assume oggi la forma del compito di alfabetizzazione digitale, e d'altra parte di cura perché il modello sociale improntato alla solidarietà e alla cura di chi è più fragile, connesso all'annuncio della gratuità dell'amore per il quale gli esseri umani condividono il dono della vita, raggiunga gli esseri umani nelle reali condizioni in cui si trovano.

Non consentire che una rivoluzione di tale portata sia gestita da forze economiche e politiche guidate esclusivamente da logiche di profitto è una sfida importante del presente: ecco perché l'acquisizione di competenza e la capacità di gestire il potenziale di questa tecnologia in modo etico e improntato alla coerenza con l'annuncio evangelico diventa una via concreta e significativa proposta nelle chiese e dalle chiese per il mondo.

Gli utilizzi potenzialmente efficaci di IA nella vita delle chiese sono molteplici: dalla gestione dei dati amministrativi, all'organizzazione del materiale didattico, alla costituzione di database intelligenti e condivisi di documenti legati alla ricerca biblica e teologica, alla formazione, alla produzione di percorsi formativi o alla gestione del lavoro diaconale. L'IA è già entrata di fatto nella gestione dei siti e nelle interfacce digitali con cui le chiese offrono informazioni all'esterno e sono presenti sulle piattaforme social, e nei vari settori delle attività gestite direttamente o indirettamente dalle Chiese.

Non si tratta dunque solo del sorriso o della preoccupazione che possono suscitare i robot dispensatori di citazioni bibliche o di "benedizioni", i supporti IA per la preghiera proposti dall'industria e dal marketing religioso in diverse realtà, ma anche di un reale supporto al lavoro che nelle Chiese e dalle Chiese viene svolto. In questo senso la capacità di offrire una testimonianza di utilizzo consapevole, efficace ed etico diventa centrale per le Chiese che possono mostrare come la preparazione anche tecnica all'interno di un orizzonte teologico sia a servizio delle relazioni e non un pericoloso sostitutivo.

Occorre oggi tenere conto e studiare il fenomeno emergente del sempre più massiccio rivolgersi alle IA per rispondere al bisogno di supporto spirituale, di assistenza e formazione, e a tutta

una serie di prospettive inedite proprio dal punto di vista dell'esperienza religiosa: la ricchezza delle tradizioni da cui proviene il Cristianesimo nella sua espressione riformata consente di leggere queste realtà mantenendo punti di riferimento essenziali, offrendo interpretazioni, ma anche formulando creativamente possibilità di utilizzo coerenti con l'annuncio evangelico.

Conclusioni.

Il presente Documento rappresenta la prima presa di contatto del lavoro della Commissione con il tema dell'Intelligenza Artificiale; come si è detto in apertura la scelta di offrire uno strumento che consenta un approccio chiarificatore e critico alla trasformazione che stiamo sperimentando ha il significato di inaugurare un possibile percorso. La complessità, la pervasività, l'esponenziale sviluppo di possibilità e problemi richiede un continuo aggiornamento delle conoscenze e un costante impegno di pensiero, ragioni che hanno condotto la Commissione a scegliere di non affrontare immediatamente un tema specifico, ma di offrire uno sguardo generale sull'argomento.

Possiamo riassumere brevemente come conclusione che tuttavia non chiude, ma al contrario inaugura un cammino, alcuni punti essenziali:

- non avere timore dell'innovazione significa imparare a utilizzare al meglio l'IA, il suo potenziale a servizio della vita delle comunità: apprendere a portare contenuti, informazioni, materiale formativo e informativo in rete è un primo livello di utilizzo in cui alcune Chiese si sono già impegnate;

- mantenere lucidamente la consapevolezza da trasmettere alle giovani generazioni che per quanto sofisticata, pervasiva e trasformativa l'IA è portatrice di opportunità solo per chi si è formato al pensiero critico e alla capacità di buon utilizzo, e invece portatrice di rischi per chi viene condizionato dai suoi effetti retroattivi, è compito della comunità credente e pensante insieme a tutti e tutte gli/le umani/e;

- mantenere vigile l'attenzione sui processi in corso a livello globale, con una chiara visione del valore dei principi che faticosamente l'Occidente ha affermato nel suo modello di vita pluralistico e democratico, lavorare con ogni altra forza che intenda indirizzare eticamente verso il bene comune il potenziale di questa rivoluzione tecnologica è un ulteriore compito al quale le Chiese sono chiamate nel porre in atto un annuncio di libertà per tutte e tutti che sia attento anche ai pericoli che gli esseri umani più fragili corrono.

Si tratta di accogliere la responsabilità di stare nella realtà che cambia, di vivere in un mondo che va verso scenari solo immaginati, e forse neppure del tutto immaginabili, senza perdere la capacità di interrogare e interrogarsi contribuendo al cammino degli esseri umani verso l'avvenire.

Glossario.

- **Addestramento (Training):** *La fase cruciale in cui un modello di IA apprende, analizzando milioni di esempi per scoprire schemi e associazioni. Durante questo processo, il modello regola i suoi **parametri** interni per migliorare l'accuratezza delle sue risposte.*
- **Algoetica:** *Il campo di studio che analizza le implicazioni etiche degli algoritmi e come possano essere governati da principi morali.*
- **Algoritmo:** *è una sequenza di operazioni, di istruzioni strutturate, che consente di risolvere un certo tipo di quesito o di risolvere un determinato tipo di problemi.*
- **Allucinazioni:** *casi in cui l'IA offre risposte inventate, con informazioni non pertinenti, imprecise o scorrette. Al momento non è del tutto eliminabile.*
- **Apprendimento Supervisionato (Supervised Learning):** *E' un addestramento in cui il modello impara da un dataset, cioè con la risposta corretta.*
- **Apprendimento Non Supervisionato (Unsupervised Learning):** *il modello viene addestrato con dati non organizzati. Lo scopo è trovare da solo schemi, strutture nascoste e cluster.*
- **Bias:** *pregiudizio, distorsione delle risposte date da un sistema IA a causa della qualità dei dati di addestramento. Può riflettere pregiudizi e generare disuguaglianza.*
- **Big Data:** *Volumi enormi e complessi di dati che non possono essere gestiti con gli strumenti tradizionali. L'IA è la tecnologia chiave per analizzare i Big Data e scoprirne valore e tendenze nascoste.*
- **Black Box (Scatola Nera):** *Un termine usato per descrivere un modello di IA il cui processo decisionale interno è opaco o incomprensibile per gli esseri umani. Non è chiaro "come" o "perché" il modello sia giunto a una determinata conclusione.*

- **Chatbot:** programma informatico che imita la conversazione con un soggetto umano, disponibile su app mobile o sito web.
- **ChatGPT:** Un modello linguistico conversazionale sviluppato da OpenAI, basato sull'architettura GPT, diventato uno degli esempi più noti di IA generativa.
- **Deepfake:** Una tecnica che utilizza l'IA per creare o manipolare video o audio in modo estremamente realistico, ad esempio sovrapponendo il volto di una persona sul corpo di un'altra. Questa tecnologia comporta seri rischi legati alla disinformazione e alle frodi.
- **Deep Learning:** campo di ricerca basato sulle reti neurali.
- **Intelligenza Artificiale:** capacità delle macchine di svolgere compiti, apprendere autonomamente, risolvere problemi con risultati che appaiono simili all'agire umano.
- **Intelligenza Artificiale Generativa:** tipo di IA in grado di produrre testi, immagini, video, musica e altri media secondo le richieste ricevute. E' alla base del funzionamento di ChatGPT, Gemini, MetaAI, Grok, Copilot, Deepseek, Claude.
- **LLM, Large Language Model:** programmi che riconoscono e producono testi sulla base di una quantità elevatissima di dati di vario tipo, testi video, immagini ecc...La scelta dei dati di addestramento determina la qualità del modello.
- **Machine Learning:** branca di studio che intende portare le macchine ad apprendere imitando il modo umano, senza istruzioni definite, generalizzando da casi, individuando schemi e lavorando su alta probabilità nelle sue previsioni.
- **Natural Language Processing (NLP):** L'Elaborazione del Linguaggio Naturale è la branca dell'IA che si occupa di permettere ai computer di comprendere, interpretare e generare il linguaggio umano.

- **Open-AI:** laboratorio di ricerca sulle IA costituito inizialmente da società no-profit e di cui fa oggi parte anche una società sussidiaria a scopo di lucro (OpenAI Gruppo PBC).
- **Prompt:** istruzioni fornite a un sistema IA. La capacità di fornirle efficacemente è un ambito sempre più importante del lavoro.
- **Rete Neurale (Artificial Neural Network):** modello matematico costituito da algoritmi che lavora imitando i circuiti neuronali biologici. Le reti convoluzionali elaborano immagini.
- **Token:** L'unità minima di testo che un modello di linguaggio elabora, che può corrispondere a una parola, una parte di essa o un singolo carattere. Il processo di suddivisione del testo in queste unità è chiamato **tokenizzazione**.
- **Transformer:** Un'architettura di rete neurale introdotta nel 2017 che ha rivoluzionato il campo dell'NLP. Basata sul meccanismo di self-attention, ha permesso di creare modelli molto più potenti ed efficienti, come BERT e la serie GPT.
- **Turing Test:** Un test proposto da Alan Turing per determinare se una macchina sia in grado di esibire un comportamento intelligente indistinguibile da quello di un essere umano.

.