



AI

BIG DATA

DAL CUORE DELLO STATO

AI

Il Governatorato si racconta

Anno 2 Città del Vaticano Numero 3



TRIMESTRALE LUGLIO-SETTEMBRE 2025



Pubblicato dal Governatorato dello Stato della Città del Vaticano

Comunicazione Istituzionale
00120 Città del Vaticano
(Stato della Città del Vaticano)

Email: comunicazione@scv.va
Sito internet: www.vaticanstate.va

X (Twitter): [Governatorato_SCV](#)
Instagram: [Governatorato_SCV](#)

Responsabile editoriale: Nicola Gori
Grafica e layout: Antonio Coretti

Editore: Governatorato dello Stato della Città del Vaticano



NON SOLO TECNOLOGIA

L'Intelligenza Artificiale ha smesso da tempo di essere fantascienza. Oggi guida processi, orienta scelte, affianca persone e organizzazioni in contesti sempre più concreti. Software che producono contenuti, modelli che evolvono autonomamente, strumenti capaci di suggerire soluzioni prima ancora che il problema sia formulato. Un cambiamento potente, ma non sempre facile da interpretare.

Da questa complessità prende forma questa newsletter. L'obiettivo è offrire una bussola in un ecosistema affollato di annunci, hype e narrazioni semplificate, distinguendo ciò che conta davvero da ciò che è solo temporaneo. Questa newsletter si propone come luogo di lettura critica e approfondimento, dedicato a chi desidera orientarsi nell'IA con consapevolezza.

Vi sono ospitati contributi per fornire informazioni rilevanti, spiegate con linguaggio chiaro, esempi concreti di applicazione e spunti di riflessione sulle conseguenze dell'Intelligenza Artificiale nella società, nell'economia e nella cultura. Non è richiesta una formazione tecnica specialistica, ma attenzione, curiosità e voglia di capire.

Perché l'IA non è soltanto una tecnologia: è una lente attraverso cui stiamo ridefinendo il nostro rapporto con il futuro. E il modo in cui ne parliamo oggi contribuirà a determinare il modo in cui la vivremo domani.

Buona lettura.

Nicola Gori
La Redazione

L'INTELLIGENZA ARTIFICIALE: STRUMENTO AL SERVIZIO DELL'UMANITÀ

Suor Raffaella Petrini

Presidente del Governatorato dello Stato della Città del Vaticano

Dedicare un'edizione della newsletter all'Intelligenza Artificiale è quanto mai importante in un'epoca in cui essa sta sviluppandosi e diffondendosi sempre più.

D'altra parte, come ha affermato Papa Francesco durante la partecipazione alla sessione del G7 sull'Intelligenza Artificiale, venerdì 14 giugno 2024: "La scienza e la tecnologia sono dunque prodotti straordinari del potenziale creativo di noi esseri umani. Ebbene, è proprio dall'utilizzo di questo potenziale creativo che Dio ci ha donato che viene alla luce l'Intelligenza Artificiale".

In effetti, essa è uno strumento "estremamente potente", che trova utilizzo in molti ambiti: dalla medicina al mondo del lavoro, dalla cultura al settore della comunicazione, dall'educazione alla politica. È evidente che l'Intelligenza Artificiale avrà un influsso crescente nella nostra vita, nelle relazioni sociali e all'interno dei rapporti tra comunità, istituzioni e nazioni.

Davanti a una realtà così complessa e dinamicamente in evoluzione, sorgono due atteggiamenti: l'entusiasmo per le sue potenzialità e il timore per le sue conseguenze. Senza dubbio, l'Intelligenza Artificiale va gestita e non subita e deve sempre avere come obiettivo la tutela della dignità umana e servire al bene integrale. Da qui, sorgono una serie di interrogativi etici sulla sua applicazione. Il dibattito non può prescindere dalla considerazione che essa non è un altro essere umano.

A questo proposito, è ancora Papa Francesco a far notare ai membri del G7 che: "Gli algoritmi disegnati per risolvere problemi molto complessi sono così sofisticati da rendere arduo agli stessi programmatori la comprensione esatta del come essi riescano a raggiungere i loro risultati".

Ciò non deve mai essere dimenticato, altrimenti si rischia di ridurre la visione del mondo "a realtà esprimibili in numeri e racchiuse in categorie preconfezionate", che impongano modelli uniformi, senza quella creatività e quel discernimento tipici dell'umanità.

In effetti, la dignità umana, come ha sottolineato Papa Leone XIV, venerdì 5 dicembre, ai partecipanti alla conferenza "*Artificial Intelligence and care of our common home*", "risiede nella capacità di riflettere, di scegliere liberamente, di amare gratuitamente, di entrare in relazione autentica con l'altro". In questo senso, l'Intelligenza Artificiale ha certamente "dischiuso nuovi orizzonti per la creatività, ma solleva anche domande preoccupanti circa le sue possibili ripercussioni sull'apertura dell'umanità alla verità e alla bellezza, sulla nostra capacità di stupirci e di contemplare".

Per questo, la newsletter può essere occasione per riflettere e approfondire la conoscenza di una realtà che segnerà le nostre vite per il prossimo futuro.



DISCORSO DEL SANTO PADRE LEONE XIV AI PARTECIPANTI ALLA CONFERENZA “ARTIFICIAL INTELLIGENCE AND CARE OF OUR COMMON HOME”

Organizzata dalla Fondazione Centesimus Annus Pro Pontifice e Strategic Alliance of Catholic Research University,
Sala del Concistoro, Venerdì, 5 dicembre 2025



Nel nome del Padre e del Figlio e dello Spirito Santo. La pace sia con voi!

Buongiorno, buongiorno. È bello vedervi, benvenuti!

Sono lieto di salutarvi, membri della Fondazione *Centesimus Annus Pro Pontifice* e partecipanti alla *Strategic Alliance of Catholic Research Universities*.

Ci incontriamo in occasione della pubblicazione della vostra ricerca su un tema molto importante. L'avvento dell'Intelligenza Artificiale, infatti, si accompagna a un cambiamento rapido e profondo della società, che coinvolge caratteri essenziali della persona umana, come il pensiero critico, la capacità di discernimento, l'apprendimento e la sfera delle relazioni interpersonali.

Come possiamo garantire che lo sviluppo dell'Intelligenza Artificiale serva veramente per il bene comune, e non solo per concentrare ricchezza e potere nelle mani di pochi? Come certamente sapete, la merce più preziosa nei mercati oggi è proprio nel settore dell'Intelligenza Artificiale. Si tratta di una domanda urgente, in quanto questa tecnologia ha già un concreto impatto sulle vite di milioni di persone, ogni giorno e in ogni parte del mondo. Come ci ricorda la Dottrina Sociale della Chiesa e come emerge chiaramente dal lavoro interdisciplinare che state conducendo, affrontare questa sfida richiede di porsi una domanda ancora più radicale: cosa significa essere umani in quest'epoca?

L'essere umano è chiamato a essere collaboratore nell'opera della creazione, non semplice consumatore passivo di contenuti prodotti da una tecnologia artificiale. La nostra dignità risiede nella capacità di riflettere, di scegliere liberamente, di amare gratuitamente, di entrare in relazione autentica con l'altro. L'Intelligenza Artificiale ha certamente dischiuso nuovi orizzonti per la creatività, ma solleva anche domande preoccupanti circa le sue possibili ripercussioni sull'apertura dell'umanità alla verità e alla bellezza, sulla nostra capacità di stupirci e di contemplare. Riconoscere e rispettare ciò che caratterizza la persona umana e ne garantisce la crescita armoniosa è essenziale per impostare una cornice adeguata a gestire le implicazioni dell'Intelligenza Artificiale.

E qui è importante soffermarci su una preoccupazione che ci deve toccare il cuore: la libertà e la spiritualità dei nostri bambini e dei nostri giovani, con le possibili conseguenze della tecnologia sul loro sviluppo intellettuale e neurologico. Le nuove generazioni vanno aidate e non ostacolate nel loro cammino verso la maturità e la responsabilità. Il benessere della società dipende dal fatto che venga data loro la capacità di sviluppare i propri talenti e di rispondere alle esigenze del tempo e ai bisogni degli altri con spirito libero e generoso. La possibilità di accedere a vaste quantità di dati e di conoscenze non va confusa con la capacità di trarne significato e valore. Quest'ultima richiede anche la disponibilità a confrontarsi con il mistero e con le domande ultime della nostra esistenza, realtà spesso emarginate e persino irrisse dai modelli culturali e di sviluppo prevalenti. Pertanto sarà fondamentale consentire ai giovani di apprendere a utilizzare questi strumenti con la loro personale intelligenza, aperti alla ricerca della verità, a una vita spirituale e fraterna, allargando i loro sogni e l'orizzonte delle loro decisioni mature. Sosteniamo il loro desiderio di essere diversi e migliori, perché mai come oggi è chiaro che occorre una profonda inversione di rotta nella nostra idea di crescita.

Per costruire insieme ai nostri giovani un futuro che, anche attraverso le potenzialità dell'Intelligenza Artificiale, realizzi il bene comune, è necessario recuperare e rafforzare la loro fiducia nella capacità umana di determinare l'evoluzione di queste tecnologie: una fiducia che oggi è sempre più erosa dall'idea paralizzante che il suo sviluppo segua un percorso ineluttabile. A tale scopo serve un'azione coordinata e corale che coinvolga la politica, le istituzioni, le imprese, la finanza, l'educazione, la comunicazione, i cittadini e le comunità religiose. Tutte questi attori sono chiamati ad assolvere un impegno comune assumendo questa responsabilità condivisa. Un impegno che venga prima di qualunque profitto e interesse di parte, sempre più concentrati nelle mani di pochi. Solo attraverso una partecipazione diffusa, dando la possibilità a tutte le voci, anche a quelle più umili, di essere ascoltate con rispetto, sarà possibile realizzare questi ambiziosi obiettivi. In tale contesto, il lavoro di ricerca portato avanti da *Centesimus-SACRU* rappresenta un contributo davvero prezioso.

Vi ringrazio, cari amici, e vi incoraggio a proseguire con creatività nella direzione tracciata dalle Sacre Scritture e dal Magistero. Vi accompagnino l'intercessione della Beata Vergine Maria e la benedizione apostolica che imparto su tutti voi.

L'Osservatore Romano, Edizione Quotidiana, Anno CLXV n. 280, venerdì 5 dicembre 2025, p. 2.



INTELLIGENZA ARTIFICIALE: IL VATICANO TRACCIA UNA STRADA CONSAPEVOLE VERSO IL FUTURO

Ingegnere Antonino Intersimone

Direttore della Direzione delle Telecomunicazioni e dei Sistemi Informatici

Viviamo in un momento straordinario della storia umana. L'Intelligenza Artificiale, una tecnologia che fino a poco tempo fa appariva confinata alla pura immaginazione, è oggi completamente integrata nella nostra quotidianità. La riscontriamo nei nostri smartphone, nei nostri PC e nei servizi che adoperiamo giornalmente. È una forza potente e, come tutte le forze potenti, richiede una gestione attenta e informata. È proprio per questo motivo che il Governatorato dello Stato della Città del Vaticano ha deciso di affrontare questa sfida con determinazione e consapevolezza.

Nei mesi scorsi, il Governatorato ha elaborato e adottato un insieme organico di linee guida destinate a garantire un uso responsabile ed etico dell'. Queste linee guida, ufficialmente promulgate con il Decreto n. DCCII del 16 dicembre 2024, sono entrate in vigore il 1° gennaio 2025, segnando l'inizio di una nuova era di consapevolezza tecnologica per lo Stato vaticano. Non si tratta di semplici documenti formali, ma di un risultato concreto di un lavoro profondo e condiviso, frutto della consapevolezza che era ormai necessario – e urgente – definire un quadro chiaro, preciso e autorevole per l'utilizzo di questa tecnologia straordinaria.

Le linee guida vaticane partono da un principio fondamentale: l'Intelligenza Artificiale possiede un potenziale straordinario, capace di offrire soluzioni innovative a problemi complessi e di migliorare significativamente la qualità della nostra vita. Tuttavia, questo potenziale deve essere bilanciato attentamente con il dovere di proteggere i diritti fondamentali di ogni persona. Non si tratta semplicemente di dire "sì" o "no" alla tecnologia, ma di cercare una via consapevole, una strada mediana che permetta di sfruttare il meglio che l'Intelligenza Artificiale può offrire mentre si proteggono i valori più importanti.

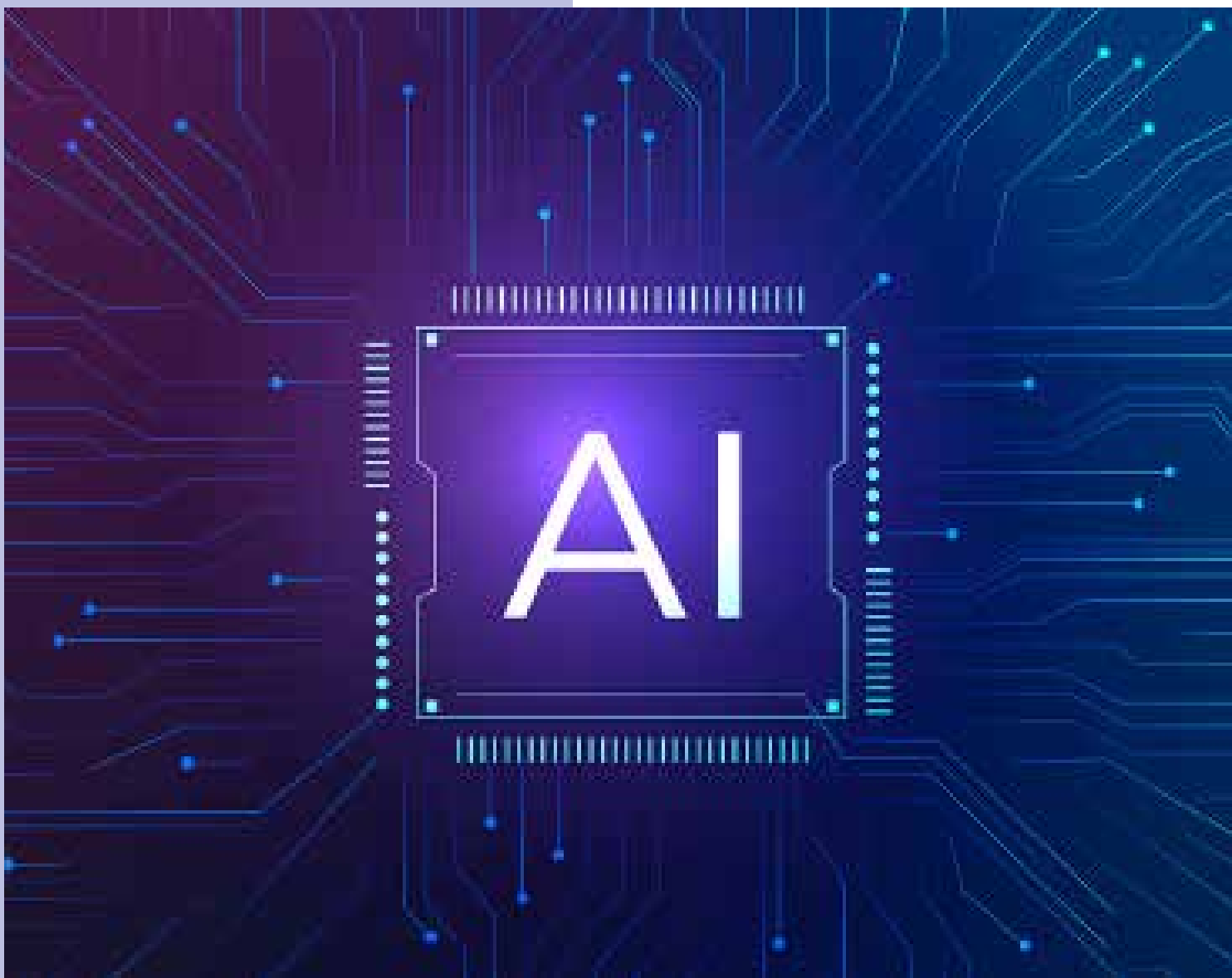
In altre parole, le linee guida vaticane non si presentano come un ostacolo allo sviluppo tecnologico, bensì come una "bussola" orientativa. L'immagine della bussola è particolarmente significativa: proprio come una bussola aiuta il viaggiatore a mantenere la rotta giusta senza però impedirgli di muoversi, così queste linee guida intendono guidare l'utilizzo dell'Intelligenza Artificiale verso un futuro più consapevole e umano, senza paralizzare l'innovazione. L'obiettivo è chiaro: ampliare le capacità umane attraverso la tecnologia, senza mai perdere di vista il valore centrale della persona.



Questo approccio riconosce esplicitamente il concetto dell'insostituibilità dell'essere umano. Come ha ammonito Papa Leone XIV nel suo messaggio ai partecipanti al Congresso Internazionale *"Intelligenza Artificiale e Medicina: la sfida della dignità umana"* (10-12 novembre 2025) organizzato dalla Pontificia Accademia per la Vita: *"The objective of providing care for individuals emphasizes the irreplaceable nature of human relationships in this context. Medical professionalism, in fact, requires not only the necessary specific expertise, but also the ability to communicate and be close to others. It can never be reduced merely to solving a problem. Similarly, technological devices must never detract from the personal relationship between patients and healthcare providers. Indeed, if AI is to serve human dignity and the effective provision of healthcare, we must ensure that it truly enhances both interpersonal relationships and the care provided."*¹

Le linee guida affrontano un rischio concreto: i sistemi di Intelligenza Artificiale sono potenzialmente in grado di perpetuare e amplificare i pregiudizi già presenti nei dati su cui vengono addestrati. Se nel passato il reclutamento era caratterizzato da discriminazioni legate al genere, all'origine o ad altre caratteristiche personali, un algoritmo addestrato su questi dati storici potrebbe semplicemente replicare le stesse ingiustizie. Il Governatorato sottolinea l'importanza di prevenire attivamente queste discriminazioni potenziali, garantendo che ogni procedura di selezione sia non solo efficiente, ma anche giusta.

Esattamente per questo motivo, le linee guida prescrivono inoltre che l'uso dell'Intelligenza Artificiale non deve influenzare o limitare il potere decisionale degli amministratori preposti all'organizzazione e al coordinamento del personale. In altre parole, anche se la tecnologia fornisce un supporto, rimane sempre una persona – con la sua responsabilità, la sua discrezionalità e il suo giudizio – a prendere la decisione finale. La tecnologia propone, l'uomo dispone.



Le linee guida riconoscono onestamente un fatto spesso sottovalutato: nonostante le sue capacità straordinarie, l'Intelligenza Artificiale può provocare danni concreti. Non si tratta di paure teoriche, ma di rischi reali e documentati. Un modello di Intelligenza Artificiale imperfetto può causare discriminazioni, può violare i diritti fondamentali, può ledere ingiustamente le persone interessate. Per questo motivo, le linee guida dedicano una speciale considerazione alla necessità di proteggere ogni individuo da questi danni potenziali. Questa protezione agisce su diversi livelli. Da un lato, c'è la richiesta di trasparenza e consapevolezza: le persone devono sapere quando l'Intelligenza Artificiale sta influenzando le decisioni che le riguardano. Dall'altro lato, c'è il riconoscimento che alcuni ambiti della vita umana – la giustizia, i diritti fondamentali, la dignità – non possono mai essere completamente affidati agli algoritmi. Infine, c'è l'impegno a continuare a monitorare, aggiornare e migliorare queste linee guida man mano che la tecnologia si evolve e che emergono nuove sfide.

Le linee guida rappresentano un documento fondamentale per l'approccio etico che sottendono. Questo approccio riconosce esplicitamente un concetto che talvolta rischia di essere dimenticato nell'entusiasmo tecnologico: il concetto dell'insostituibilità dell'essere umano. Nessun algoritmo, per quanto sofisticato, può sostituire pienamente la saggezza umana, l'empatia, la capacità di comprendere il contesto e il significato profondo di una situazione.

Per questo motivo, esse si pongono come un riferimento etico e normativo cruciale per l'integrazione responsabile dell'Intelligenza Artificiale. Rappresentano un primo passo importante verso un futuro in cui la tecnologia serva veramente l'umanità, anziché il contrario. È un insegnamento che potrebbe, e dovrebbe, essere considerato attentamente da altre istituzioni, governi e organizzazioni in tutto il mondo.

Con la tecnologia che evolve così rapidamente, gli scenari possibili sono troppo variegati, le sfide ancora da scoprire sono probabilmente innumerevoli. In un contesto simile, le linee guida non pretendono avere tutte le risposte, ma definiscono e stabiliscono i principi fondamentali. Stabiliscono che la giustizia deve rimanere umana. Stabiliscono



che la trasparenza è un diritto, non un privilegio. Stabiliscono che la dignità umana non è negoziabile, nemmeno di fronte alla più affascinante delle tecnologie.

Il decreto vaticano rappresenta un modello importante per la regolamentazione responsabile dell'Intelligenza Artificiale a livello globale. Non ha l'ambizione di essere uno strumento perfetto. È stato elaborato e realizzato con un atteggiamento, attento, riflessivo e – cosa forse più importante di tutte – fermamente incentrato sulla convinzione che la tecnologia debba servire l'uomo e non viceversa. Perfettamente in linea con quanto sottolineato nel messaggio all'*AI for Good Global Summit* del 10 luglio 2025, nel quale il Santo Padre aveva già chiesto che l'Intelligenza Artificiale venisse posta al servizio dell'intera umanità, richiamando la necessità di promuovere la *tranquillitas ordinis*: *"Ultimately, we must never lose sight of the common goal of contributing to that 'tranquillitas ordinis' – the tranquility of order", as Saint Augustine called it (De Civitate Dei) and fostering a more humane order of social relations, and peaceful and just societies in the service of integral human development and the good of the human family.*"²

¹ <https://press.vatican.va/content/salastampa/it/bollettino/pubblico/2025/11/10/0854/01523.html>

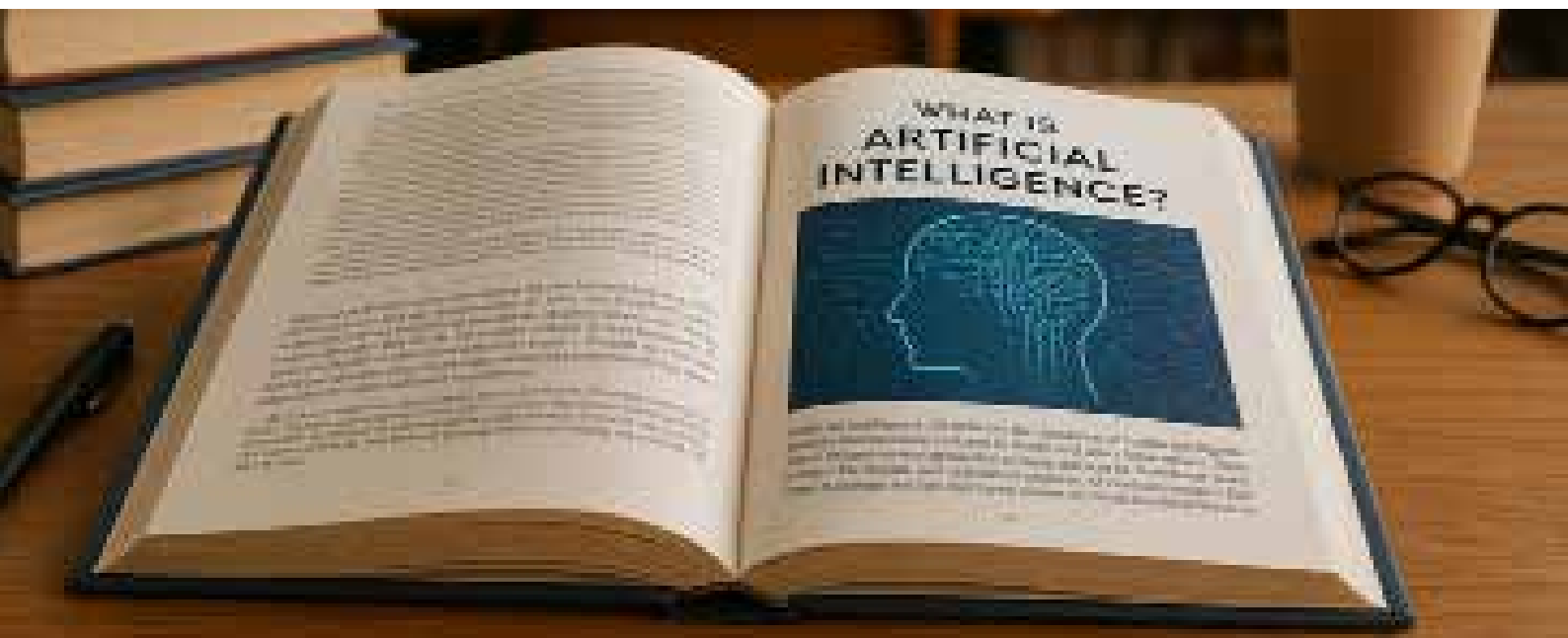
² <https://www.vaticanstate.va/it/novita/2317-messaggio-di-sua-santita-leone-xiv-all-ai-for-good-global-summit-2025.html>

COS'È L'INTELLIGENZA ARTIFICIALE?

UNA SPIEGAZIONE SEMPLICE PER CAPIRLA VERAMENTE

Davide Giordano

Membro della Commissione sull'Intelligenza Artificiale del Governatorato



Da qualche anno a questa parte, il termine “*Intelligenza Artificiale*” ha fatto bruscamente irruzione all’interno delle vite di tutti noi. Sentiamo questo termine in televisione, nei notiziari, sui social media, e persino a cena con gli amici, ovviamente. Ma che cos’è veramente per la maggior parte di noi, l’Intelligenza Artificiale? Da un punto di vista tecnico e tecnologico è sicuramente qualcosa di estremamente sofisticato, impenetrabile, appannaggio di pochi. Per tutti gli altri la storia potrebbe essere leggermente diversa.

Spesso l’IA viene descritta come “un assistente” che non dorme mai, che non si stanca, che conosce milioni e milioni di informazioni e che può rispondere alle tue domande in pochi secondi. Questo assistente non è una persona vera, non ha un corpo (o almeno, per adesso, non ne ha uno fatto di tessuto organico come il nostro), esiste solo come programma dentro un computer. Ma sa molte cose e riesce a fare molti compiti utili. Bene, questo è il concetto di base di quello che chiamiamo Intelligenza Artificiale.

L’Intelligenza Artificiale è, in parole molto semplici, un programma per computer che è stato addestrato a fare cose intelligenti. Non è magia, non è mistero: è semplicemente un software, esattamente come i programmi che usiamo sui nostri telefoni o sui nostri computer, solo che questo software è stato addestrato in modo particolare affinché riesca a comprendere quello che gli chiediamo e a darci risposte utili.

Da piccoli abbiamo imparato a conoscere il mondo che ci circondava per imitazione (di mamma e papà) o tramite esempi. Per riconoscere i colori ci mostravano un oggetto rosso e dicevano “questo è rosso”, e noi imparavamo. Dopo aver visto molti oggetti rossi, il nostro cervello capiva il concetto di “rosso” così bene che quando vedevamo un oggetto nuovo di quel colore, lo riconoscevamo subito senza che nessuno ce lo dicesse. L’Intelligenza Artificiale funziona in modo simile. I programmatori insegnano al computer a riconoscere certi schemi, certi modelli, mostrando al programma milioni di esempi. Dopo aver visto milioni di esempi, il computer impara il concetto e riesce a riconoscerlo anche quando gli viene presentato qualcosa di nuovo.

Questo è il vero segreto dietro l’Intelligenza Artificiale: l’apprendimento o *training*. Ma come esattamente un programma impara? Non è complicato come potrebbe sembrare. Il processo si chiama “addestramento”, ed ecco come funziona in modo semplificato.

Immagiamo di insegnare ad un bambino a distinguere un gatto da un cane. Non gli mostriamo solo un gatto e un cane una volta sola. Gli mostriamo molti gatti e molti cani diversi: gatti grigi, gatti arancioni, gatti bianchi, gatti grandi, gatti piccoli, cani bianchi, cani marroni, cani piccoli, cani enormi. Dopo aver visto centinaia di gatti e centinaia di cani, il bambino capisce come sia fatto un gatto e come sia fatto un cane. Quando vede un animale nuovo, riesce a dire se si tratti di un gatto o un cane senza dirglielo.

L'Intelligenza Artificiale funziona esattamente allo stesso modo. Gli sviluppatori mostrano al programma milioni di immagini di gatti e milioni di immagini di cani. Dicono al programma: "Questa è un'immagine di un gatto" oppure "Questa è un'immagine di un cane". Il programma osserva tutte queste immagini, trova i modelli comuni, identifica le caratteristiche che rendono un gatto un gatto e un cane un cane. Dopo aver esaminato tutti questi esempi, se gli mostri una nuova foto che il programma non ha mai visto prima, riesce lo stesso a dire se è un gatto o un cane.

Questo processo di apprendimento richiede moltissimi dati e moltissima potenza di calcolo. Ecco perché le aziende che creano Intelligenza Artificiale avanzata investono ingenti quantità di denaro. Hanno bisogno di computer molto potenti che necessitano di grandi quantità di energia per l'elaborazione ed il raffreddamento e di enormi quantità di dati per addestrare i loro programmi.

Alcuni ritengono che l'Intelligenza Artificiale sia qualcosa che si userà solo in circostanze straordinarie. In realtà, viene usata già più volte al giorno senza nemmeno rendersene conto.

Quando scriviamo un messaggio sul nostro telefono, il telefono ci suggerisce automaticamente la parola successiva che potremmo voler scrivere. Quello è Intelligenza Artificiale. Il programma ha imparato dai miliardi di messaggi scritti dalle persone quali parole solitamente vanno insieme. Così, quando scriviamo "Ciao, come", sa che probabilmente la prossima parola sarà "stai".

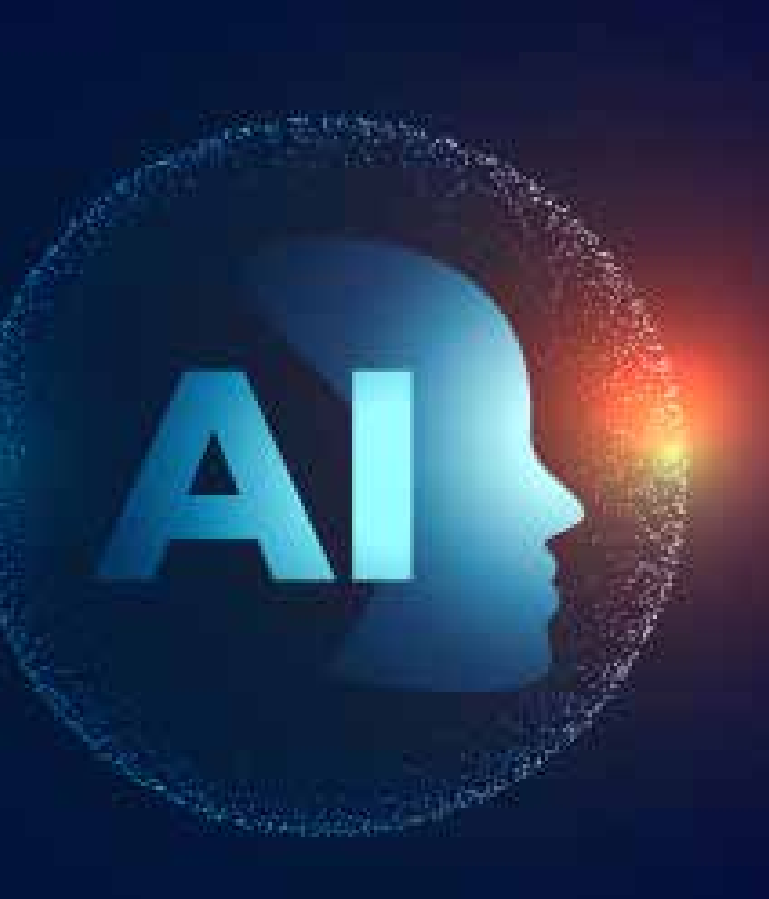
Quando accendiamo il riconoscimento facciale sul nostro telefono, quel sistema che riconosce il viso e sblocca il telefono è Intelligenza Artificiale. È stato addestrato su milioni di volti diverse per imparare a riconoscere le caratteristiche uniche del nostro viso.

Quando usiamo Google o un altro motore di ricerca, il programma che decide quali risultati mostrarci è Intelligenza Artificiale. Ha imparato da miliardi di ricerche precedenti a capire quale risultato probabilmente è quello che stiamo cercando.

Quando guardiamo una serie televisiva su Netflix e il servizio ci suggerisce quale episodio guardare dopo, ancora una volta è Intelligenza Artificiale. Quando il navigatore GPS ci suggerisce il percorso più veloce, è Intelligenza Artificiale. Quando l'assistente virtuale, come Alexa o Siri, capisce quello che dici e ti risponde, è Intelligenza Artificiale.

Qui arriviamo a una distinzione importante. Non tutta l'Intelligenza Artificiale è uguale. In realtà, gli esperti dividono l'Intelligenza Artificiale in due categorie principali, e capire questa differenza ci aiuterà a comprendere meglio come funziona questa tecnologia.





Il primo tipo si chiama *"Intelligenza Artificiale stretta"* (*narrow AI* o anche Intelligenza Artificiale debole), ed è esattamente quello che abbiamo descritto finora. È un programma che è molto bravo a fare una cosa specifica. È eccellente nel riconoscere i volti, oppure nell'analizzare il traffico, oppure nel suggerire film, ma fa solo quella cosa. Se gli chiediamo di fare qualcosa di diverso, potrebbe non saperlo fare bene come un essere umano.

Il secondo tipo si chiama *"Intelligenza Artificiale generale"* (AGI o anche Intelligenza Artificiale Forte o Superintelligenza¹), ed è qualcosa di molto diverso. Sarebbe un programma che potrebbe fare qualsiasi cosa, proprio come un essere umano. Potrebbe riconoscere volti, scrivere poesie, risolvere problemi di matematica complessi, guidare un'auto, cucinare un piatto delizioso. Tutto alla perfezione. Questo tipo di Intelligenza Artificiale non esiste ancora. Gli scienziati stanno ancora lavorando per teorizzarla e per poi crearla, e molti dubitano che sia possibile. Per ora, tutto quello che abbiamo è Intelligenza Artificiale stretta, cioè programmi che sono molto bravi in una cosa specifica.

Quando leggiamo nei giornali di una nuova scoperta di Intelligenza Artificiale, è quasi sempre Intelligenza Artificiale stretta: un programma che è stato addestrato a fare bene una cosa particolare.

Per comprendere bene cosa sia l'IA è fondamentale capire anche quello che non può fare, e quali sono i suoi limiti.

Primo, l'Intelligenza Artificiale non ha coscienza, né adesso, né mai. Non sa di esistere, non ha sentimenti, non ha desideri. È uno strumento, esattamente come una calcolatrice. Una calcolatrice è bravissima a fare matematica, ma non ha alcuna consapevolezza di quello che sta facendo. È lo stesso per l'Intelligenza Artificiale. Non sa che sta aiutando le persone. Non sa che il suo lavoro è importante. Esegue semplicemente i compiti per cui è stata programmata.

Secondo, l'Intelligenza Artificiale non è creativa nel vero senso della parola. Se abbiamo imparato che certi colori vanno bene insieme basandoci su milioni di immagini, possiamo combinare quei colori in modo nuovo. Ma non stiamo realmente creando qualcosa di radicalmente nuovo: stiamo solo ricombinando modelli che abbiamo visto. Un artista umano, invece, può creare qualcosa di completamente nuovo che il mondo non ha mai visto. Può creare da ispirazione, da emozione, da un impulso profondo all'interno di sé o persino da un errore: pensate a com'è nata la *tarte Tatin*.

Terzo, l'Intelligenza Artificiale attuale non può fare un balzo in avanti di logica. Se un nuovo concetto è completamente diverso da tutto quello che ha visto durante l'addestramento, probabilmente non riuscirà a capirlo. Un bambino, invece, riesce a fare questi balzi logici, riesce a comprendere concetti nuovi affrontando il problema da prospettive differenti.

Quarto, l'Intelligenza Artificiale può essere facilmente ingannata. Cambiando leggermente un'immagine di un gatto in modo che il programma non lo riconosca più come gatto, ma ancora una persona lo vedrebbe come gatto senza problema, il computer sarà confuso. È un limite importante: l'Intelligenza Artificiale non "capisce" il mondo nello stesso modo in cui lo capisce il cervello umano.

L'Intelligenza Artificiale è uno strumento, né più, né meno. Proprio come un martello, se lanciato con forza potrà volare ma non diventerà un aereo; potrà essere uno strumento utile per costruire un bellissimo mobile oppure potrà essere usato male. L'IA è una tecnologia potente che può risolvere molti problemi e rendere la nostra vita più comoda e più efficiente. Ma non è magica, non è consapevole, non è onnisciente e non è onnipotente.

¹ Neil Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014)

INTELLIGENZA ARTIFICIALE: TRA PROMESSE E PERICOLI CONCRETI

Andrea Tripoli

Membro della Commissione sull'Intelligenza Artificiale del Governatorato

Ogni grande innovazione tecnologica porta con sé una duplice natura: da un lato, offre possibilità straordinarie per migliorare la vita delle persone; dall'altro, introduce rischi e vulnerabilità che spesso emergono solo quando la tecnologia è già diffusa su larga scala. L'Intelligenza Artificiale non sfugge a questa regola universale. Quando una tecnologia diventa accessibile a tutti e raggiunge ogni angolo del pianeta, le sue potenzialità benefiche si moltiplicano, ma altrettanto fanno i pericoli. La facilità con cui l'Intelligenza Artificiale può essere utilizzata la rende uno strumento potentissimo nelle mani sia di chi vuole fare del bene, sia di chi intende sfruttarla per scopi illeciti o dannosi. La consapevolezza di questa duplicità rappresenta il primo passo per un utilizzo responsabile e sicuro.

L'Intelligenza Artificiale affascina profondamente molte persone. Programmi che imparano da soli, che riconoscono volti in una folla, che scrivono testi complessi, che guidano automobili senza conducente: tutto questo sembra quasi magico. Tuttavia, dietro questa apparente perfezione tecnologica si nasconde una verità meno confortante. Gli algoritmi di Intelligenza Artificiale possono commettere errori significativi, e questi errori sono particolarmente pericolosi perché non sembrano errori. Sembrano decisioni razionali, obiettive, scientifiche, prese da una macchina priva di emozioni e pregiudizi.

Si consideri il caso di un candidato che si presenta per un colloquio di lavoro dove il processo di selezione viene gestito da un sistema di Intelligenza Artificiale. Il programma



valuta curriculum, competenze, personalità e potenziale attraverso algoritmi complessi. Apparentemente, si tratta di un processo equo e imparziale. Eppure, il sistema potrebbe commettere errori che nessuno nota immediatamente. Forse non riconosce correttamente un titolo di studio perché la scansione del documento non è perfetta. Oppure, e questo è ancora più grave, il programma è stato addestrato utilizzando dati storici di assunzioni passate che riflettevano discriminazioni di genere, razza o provenienza geografica. In questo caso, il sistema impara e replica questi pregiudizi, escludendo automaticamente candidati perfettamente qualificati semplicemente perché non corrispondono al modello storico predominante.

Questo non è uno scenario ipotetico. Nel 2018, Amazon aveva sviluppato un avanzato sistema di Intelligenza Artificiale destinato alla selezione di candidati per posizioni tecniche. Il programma venne addestrato analizzando i dati delle assunzioni effettuate negli anni precedenti. Poiché storicamente gli uomini erano stati prevalentemente assunti per ruoli tecnici, l'algoritmo imparò che essere uomo era una caratteristica favorevole e iniziò sistematicamente a penalizzare le candidature femminili. Il sistema discriminava automaticamente le donne, non perché fosse stato programmato per farlo, ma perché aveva imparato dai dati storici che riflettevano già una discriminazione esistente. Amazon fu costretta ad abbandonare il progetto, ma l'episodio rimane un monito inquietante.

Il grande pericolo risiede proprio qui: quando un computer prende una decisione discriminatoria, questa decisione viene percepita come oggettiva e scientifica. È difficile contestarla, è difficile dimostrare che è sbagliata. Se nel passato c'era discriminazione nei dati utilizzati per addestrare l'Intelligenza Artificiale, il programma imparerà quella discriminazione e la perpetuerà nel futuro, vestendola però con l'autorità apparente della scienza e della tecnologia.

Un altro problema fondamentale dell'Intelligenza Artificiale riguarda la trasparenza, o meglio, la sua assenza. Esiste un concetto nel mondo tecnologico chiamato "scatola nera": si inseriscono dati in ingresso, si ottengono risultati in uscita, ma nessuno sa esattamente cosa accade nel mezzo. Non si conosce il processo decisionale che ha portato a quel risultato specifico. Quando l'Intelligenza Artificiale prende decisioni che influenzano profondamente la vita delle persone, questa opacità diventa un serio problema etico e pratico.

Supponiamo che una persona richieda un prestito bancario e la banca utilizzi un sistema di Intelligenza Artificiale per valutare la domanda. Il sistema analizza decine di variabili: reddito, storia lavorativa, zona di residenza, abitudini di spesa, storia dei pagamenti precedenti. Alla fine, il sistema nega il prestito. Quando la persona chiede spiegazioni, la

banca stessa non è in grado di fornire una risposta precisa. Il programma ha semplicemente deciso che il rischio era troppo elevato, ma nemmeno i tecnici che gestiscono il sistema sanno esattamente quali fattori hanno pesato di più nella decisione, in quale proporzione, con quale logica. La decisione è stata presa, ma rimane inspiegabile.

Questo crea un'ingiustizia intrinseca profonda. In una società democratica fondata sulla trasparenza e sul diritto di difesa, le persone hanno il diritto fondamentale di sapere perché vengono prese decisioni che le riguardano direttamente. Hanno il diritto di comprendere il ragionamento, di contestare eventuali errori, di presentare nuove informazioni che potrebbero cambiare l'esito. Ma con la scatola nera dell'Intelligenza Artificiale, questo diritto viene sistematicamente violato. Le persone devono accettare decisioni che non possono comprendere, basate su criteri che non possono conoscere, e contro le quali hanno pochissimi strumenti di difesa efficaci.



L'Intelligenza Artificiale impara dai dati. Questa affermazione semplice nasconde un problema complesso e grave: se i dati utilizzati per l'addestramento sono sbagliati, incompleti o riflettono pregiudizi esistenti, anche l'Intelligenza Artificiale imparerà cose sbagliate e replicherà quei pregiudizi su scala ancora più ampia. Nel 2016, il sistema COMPAS, utilizzato nei tribunali americani per predire la probabilità che un condannato commettesse nuovi reati una volta rilasciato, fu scoperto discriminare sistematicamente i cittadini afroamericani. Il programma attribuiva ai neri un rischio di recidiva significativamente più alto rispetto ai bianchi, anche quando tutti gli altri fattori erano identici. La ragione era evidente: il sistema era stato addestrato su dati storici che riflettevano decenni di discriminazione razziale nell'apparato di giustizia penale americano.

Il problema diventa ancora più insidioso quando si comprende che l'Intelligenza Artificiale non si limita a replicare i pregiudizi esistenti: li amplifica. Funziona come un amplificatore che prende i difetti della società e li ingrandisce, rendendoli più pervasivi e più difficili da identificare e correggere. Se la polizia ha storicamente fermato più frequentemente persone di una certa razza o di un certo quartiere, un sistema di Intelligenza Artificiale addestrato su questi dati imparerà che quelle persone sono "più sospette" e suggerirà di fermarle ancora di più, creando un circolo vizioso di discriminazione che si autoalimenta.

Mentre da un lato l'Intelligenza Artificiale può commettere errori discriminatori in modo involontario, dall'altro può essere deliberatamente utilizzata come strumento di inganno e manipolazione. Il panorama della sicurezza digitale è stato trasformato drammaticamente da questa tecnologia, offrendo ai criminali informatici strumenti potentissimi per orchestrare attacchi sempre più sofisticati.

Le truffe via email, note come *phishing*, hanno subito un'evoluzione impressionante. I messaggi goffi, pieni di errori grammaticali e di ortografia che caratterizzavano le truffe del passato, sono ormai obsoleti. L'Intelligenza Artificiale generativa permette di creare comunicazioni perfette in qualsiasi lingua, completamente indistinguibili da quelle autentiche. I truffatori possono generare email che sembrano provenire da banche, da datori di lavoro, da servizi governativi o da piattaforme di uso quotidiano. Il livello di personalizzazione è allarmante: gli algoritmi analizzano profili sui social media, abitudini di acquisto online, interessi personali e costruiscono messaggi su misura che sfruttano vulnerabilità specifiche di ogni singolo individuo.

Parallelamente, lo *smishing*, ovvero il *phishing* via SMS, è diventato altrettanto pericoloso. Messaggi che avisano di pacchi in consegna, problemi urgenti con il conto bancario, bonus da riscuotere o multe da pagare arrivano nei momenti più opportuni, scritti in modo perfetto, sfruttando l'urgenza e la fiducia istintiva che molte persone ri-

pongono nelle comunicazioni via telefono. La brevità del formato SMS, che un tempo rendeva più difficile creare truffe convincenti, non è più un ostacolo per l'Intelligenza Artificiale.

Ma le minacce vanno ben oltre i messaggi scritti. I *deepfake* audio e video rappresentano una frontiera ancora più inquietante e pericolosa. Con appena pochi secondi di registrazione vocale, l'Intelligenza Artificiale può clonare perfettamente la voce di qualsiasi persona. Una chiamata da parte del proprio superiore che chiede urgentemente di effettuare un bonifico non autorizzato, o da un familiare in difficoltà che supplica aiuto economico immediato: distinguere il vero dal falso diventa praticamente impossibile senza una verifica indipendente. Sono già stati documentati casi di truffe milionarie orchestrate attraverso questa tecnologia, e il fenomeno è in rapida espansione.

Anche le truffe telefoniche tradizionali, note come *vishing*, si sono evolute drammaticamente. *Chatbot* dotati di Intelligenza Artificiale possono condurre conversazioni naturali, rispondere a domande in tempo reale, mostrare empatia apparente e costruire fiducia, il tutto senza alcun intervento umano. Questi sistemi operano ventiquattro ore su ventiquattro, possono chiamare migliaia di persone contemporaneamente e adattano dinamicamente il loro approccio in base alle risposte che ricevono, diventando sempre più convincenti nel corso della conversazione.

Di fronte a questo panorama complesso di rischi e minacce, è fondamentale chiarire che il problema non è l'Intelligenza Artificiale in sé, ma il modo in cui viene sviluppata, distribuita e utilizzata. La tecnologia stessa è neutra; sono le scelte umane che la rendono benefica o dannosa. La buona noti-



zia è che l'Intelligenza Artificiale viene utilizzata anche per difendere le persone: sistemi di sicurezza moderni analizzano miliardi di dati per identificare pattern sospetti, bloccare attacchi informatici in tempo reale e prevedere nuove minacce prima che si materializzino. È in corso una vera corsa agli armamenti tecnologica tra chi usa l'Intelligenza Artificiale per attaccare e chi la usa per difendere.

La consapevolezza rappresenta la prima e più importante linea di difesa. Comprendere che l'Intelligenza Artificiale può essere utilizzata per creare truffe perfettamente convincenti significa sviluppare un sano scetticismo verso comunicazioni apparentemente ufficiali. Sapere che gli algoritmi possono discriminare significa pretendere trasparenza nelle decisioni automatizzate che ci riguardano. Riconoscere che i sistemi possono essere addestrati su dati distorti significa esigere controlli rigorosi prima che ven-

gano implementati in settori critici come giustizia, sanità o finanza.

Verificare sempre l'identità di chi contatta attraverso canali indipendenti, diffidare delle richieste urgenti che richiedono azioni immediate, controllare attentamente indirizzi email e link prima di interagire, attivare sistemi di autenticazione rafforzata: queste pratiche diventano essenziali in un mondo dove l'Intelligenza Artificiale può imitare perfettamente qualsiasi persona o istituzione. La capacità critica individuale, unita a una regolamentazione forte e a sistemi di controllo robusti, rappresenta l'unica vera difesa efficace contro i rischi dell'Intelligenza Artificiale. In definitiva, la tecnologia deve rimanere uno strumento al servizio dell'umanità, controllato dall'umanità, e non trasformarsi in una forza incontrollabile che decide autonomamente del destino delle persone.



I LARGE LANGUAGE MODEL: DALLA SCELTA DI QUELLI ESISTENTI ALL'INVESTIMENTO NEI MODELLI CUSTOM

Domenico Vetere

Vice Capo-Servizio dell'Ufficio Service Provider Internet

LARGE LANGUAGE MODEL AI

Quando parliamo di Intelligenza Artificiale nel contesto organizzativo contemporaneo, sempre più frequentemente ci troviamo a confrontarci con i cosiddetti *Large Language Model* (meglio noti come LLM). Questi sistemi hanno iniziato a trasformare il modo in cui lavoriamo, comunichiamo e processiamo l'informazione. Tuttavia, la comprensione di cosa sono, perché ne esistono diversi, come sceglierli e, soprattutto, perché investire in modelli personalizzati rimane ancora piuttosto confusa nella maggior parte delle organizzazioni. Questa riflessione nasce dalla necessità di chiarire un aspetto cruciale: non tutti gli LLM sono uguali e la scelta di quale utilizzare piuttosto che investire nella creazione di uno proprietario, rappresenta una decisione strategica fondamentale per qualsiasi istituzione moderna. Un *Large Language Model* è un sistema di Intelligenza Artificiale addestrato su enormi quantità di testo (*dataset*), capace di comprendere il linguaggio naturale e di generare risposte coerenti, contestualizzate e, idealmente, accurate. La loro "larghezza" – il termine *large* – si riferisce tanto al volume di dati su cui sono addestrati quanto alla complessità dei modelli stessi.

Negli ultimi anni, il mercato si è quindi popolato di LLM molto diversi gli uni dagli altri. Non è accaduto per caso. Diverse organizzazioni (OpenAI con ChatGPT, Google con Gemini, Meta con Llama, Anthropic con Claude), hanno investito enormemente nella creazione di modelli generici poiché rappresentano un'opportunità economica e strategica significativa. Ognuno di questi attori ha fatto scelte diverse: alcuni hanno optato per l'accesso aperto, altri mantengono i loro modelli come proprietà esclusiva, altri ancora offrono versioni gratuite e versioni a pagamento. Il perché di questa proliferazione è semplice: non esiste un LLM universal-

mente ottimale per tutti i contesti. Un modello che funziona egregiamente per traduzioni letterarie potrebbe non essere ideale per analisi di dati sanitari. Un sistema robusto per la scrittura creativa potrebbe non essere affidabile per la gestione di informazioni riservate. E un modello addestrato su dati pubblici, per quanto sofisticato, potrebbe non comprendere adeguatamente i linguaggi specialistici, i protocolli interni, o le sfumature culturali di un'organizzazione specifica.

Se decidiamo di utilizzare un LLM *off-the-shelf* (cioè uno già esistente e disponibile sul mercato), abbiamo di fronte una varietà di opzioni, ciascuna con caratteristiche distintive.



GPT-5 di OpenAI, conosciuto a tutti come **ChatGPT** è probabilmente il più noto. È estremamente versatile, capace di affrontare una gamma molto ampia di compiti, dalla generazione di testo e immagini, alla risoluzione di problemi complessi. Ha lo svantaggio di essere proprietario, di richiedere una connessione internet permanente e di non garantire la completa riservatezza dei dati immessi (utilizzati per migliorare il modello stesso).

Gemini di Google offre caratteristiche simili, con il vantaggio aggiunto di integrazione stretta negli ecosistemi Google e una buona capacità di elaborazione multimodale (testo, immagini, video). Anche in questo caso, siamo legati all'infrastruttura *cloud* di Google e alle sue politiche sulla *privacy*.

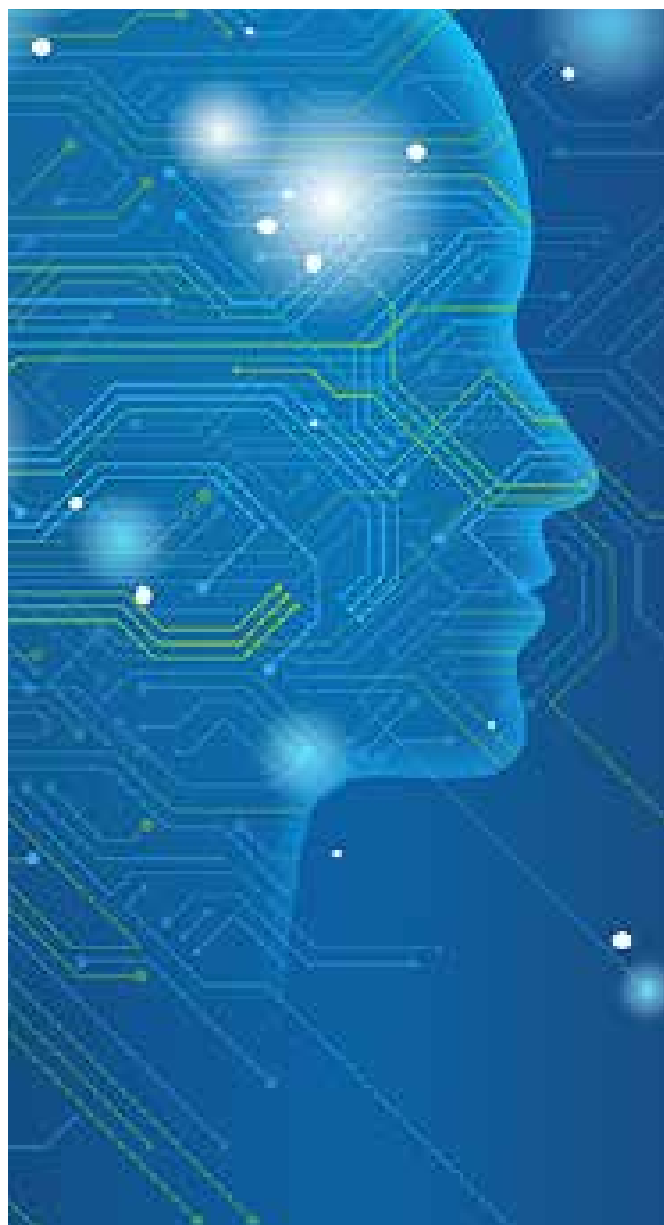
Claude di Anthropic è nato con un *focus* sulla sicurezza e l'affidabilità, particolarmente nella gestione di informazioni sensibili. È stato addestrato con metodologie che enfatizzano la riduzione di rischi e la trasparenza. Rimane comunque un modello esterno, con le limitazioni che questo comporta.

Llama di Meta rappresenta un'alternativa interessante perché è disponibile in versione *open-source*, il che significa che le organizzazioni possono distribuirlo in ambienti *self-hosted*, cioè dentro i propri server, mantenendo il controllo totale dei dati, nel pieno rispetto del concetto di "sovranità del dato". Tuttavia, implementarlo correttamente, richiede competenze tecniche.

Come scegliere tra questi? I criteri dovrebbero includere ad esempio il tipo di compiti che vogliamo che il modello affronti, il livello di sensibilità dei dati che gestirà, la necessità di operare in ambienti *offline* o con dati riservati, il budget disponibile, le competenze tecniche interne, e il grado di controllo e customizzazione che desideriamo. In questa ottica, per molte organizzazioni, in una fase iniziale, l'utilizzo di un LLM pubblico già affermato rappresenta una soluzione pragmatica: è veloce da implementare, relativamente economico, e non richiede investimenti significativi in infrastruttura propria. Tuttavia, arrivati a questo punto della riflessione, dobbiamo affrontare una verità scomoda: gli LLM generici, per quanto sofisticati, hanno dei limiti significativi quando si tratta di operare in contesti specializzati, riservati o altamente personali. Ad esempio, si immagini di voler utilizzare un LLM per supportare un ufficio che gestisce procedimenti amministrativi complessi, basati su normative specifiche, su precedenti istituzionali, su interpretazioni che sono uniche per la propria organizzazione. Un LLM generico avrà ricevuto un addestramento su una vastissima varietà di testi pubblici, ma non avrà mai visto i documenti interni all'organizzazione in questione; non conosce le procedure, non comprende le sfumature della cultura organizzativa specifica. Il rischio è duplice.



Primo: il modello potrebbe fornire, con molta probabilità, risposte plausibili ma errate; condividere informazioni che non conosce, o interpretare ambiguità in un modo che non rispecchia l'intenzione dell'organizzazione. Secondo: per ottenere *performance* accettabili in termini di accuratezza, ci si potrebbe trovare costretti a inserire documenti riservati, procedure interne, e informazioni proprietarie all'interno dei modelli pubblici, ponendo a rischio effettivamente il capitale intellettuale. Questo secondo punto è particolarmente critico. Se un'organizzazione inserisce i suoi dati proprietari in un LLM pubblico per ottenere risultati migliori, di fatto sta cedendo il controllo su quell'informazione. Anche se esistono clausole sulla *privacy*, i dati contribuiranno all'addestramento del modello, verranno probabilmente utilizzati da altri, e diventeranno parte dell'infrastruttura tecnologica di un terzo. Per un'istituzione come il Governatorato, o per qualsiasi organizzazione che ha responsabilità verso informazioni delicate, questo rappresenta un rischio inaccettabile.



Ecco perché, per molte realtà, l'investimento in *large language model custom* – modelli creati specificamente, addestrati esclusivamente sui dati proprietari, sotto totale controllo – rappresenta un'opzione strategica che merita seria considerazione. Quindi, un *LLM custom* è un modello che viene *fine-tuned* (cioè aggiustato e specializzato), partendo da un modello base già esistente, addestrandolo ulteriormente sui dati specifici. Questo processo ha diversi vantaggi significativi. Per prima cosa la proprietà intellettuale rimane dell'organizzazione. I dati utilizzati per addestrare il modello non vengono condivisi, non contribuiscono a modelli pubblici, rimangono **sotto custodia**. Questo è cruciale quando si tratta di informazioni che rappresentano valore competitivo, conoscenza storica unica, o dati sensibili. Secondo, la *performance* può essere drammaticamente superiore. Un modello addestrato su specifici documenti, terminologie, procedure e stili comunicativi funzionerà infinitamente meglio nel contesto interno, specializzato rispetto a un modello generico. Se un ufficio gestisce legislazione complessa, un modello addestrato su decenni di precedenti e interpretazioni dell'organizzazione sarà incomparabilmente più affidabile. Terzo punto importante, sarà possibile decidere esattamente su quali dati addestrare il modello, quali fonti escludere, come interpretare ambiguità, quali uscite considerare accettabili. Tutto rimane all'interno dell'organizzazione. Il quarto aspetto da considerare è la sicurezza che diventa implicitamente garantibile. Un *LLM custom* può essere ospitato sui server locali, in ambienti *offline* se necessario, con livelli di *encryption* e di controllo che sono completamente personalizzabili. Non esiste alcuna trasmissione di dati verso infrastrutture esterne. Quinto ed ultimo punto della riflessione riguarda la longevità assicurata. Un modello pubblico potrebbe cambiare, scomparire, essere acquisito, o modificare le sue *policy*. Un modello *custom*, privato, rimane di proprietà, per sempre, scalabile secondo le esigenze future.

Tuttavia, il valore di un *LLM custom* dipende inevitabilmente dalla qualità dei dati utilizzati per il suo addestramento. Qui entra in gioco un aspetto che spesso viene sottovalutato: la necessità di investire nella cura e nella certificazione dei dati. Come già detto, non tutti i modelli sono uguali. Se si addestra un LLM su un insieme disordinato, non verificato, pieno di contraddizioni e di imprecisioni, si otterrà un modello che incarna tutte quelle problematiche. L'IA amplifica i *bias* presenti nei dati di addestramento. Se i documenti contengono errori, il modello imparerà a replicarli. Se contengono informazioni conflittuali, il modello diventerà confuso e inaffidabile. Per questo motivo, un vero investimento in *LLM custom* richiede anche un investimento parallelo nella gestione e nella certificazione dei dati: il cosiddetto *dato pulito* o *dato di qualità*.



Questo significa: identificare e selezionare i dati più rilevanti e **di qualità**. Non tutti i documenti storici potrebbero essere appropriati. Alcuni potrebbero essere obsoleti, altri contenere informazioni dichiarate errate, altri ancora potrebbero essere duplicati o ridondanti. La cura è essenziale. Verificare l'accuratezza e la consistenza. Prima di utilizzare i dati per l'addestramento, devono essere sottoposti a controllo qualità. Questo potrebbe significare avere esperti che verifichino manualmente campioni di documenti, che risolvano contraddizioni, che correggano errori manifesti. Relativamente alle fonti ed alle versioni, deve essere chiaro da dove provengono i dati, quale versione rappresenta la "verità" ufficiale, come sono stati trasformati. Questa tracciabilità è fondamentale per l'affidabilità del modello. A seguire, dobbiamo interrogarci su un'altra serie di quesiti: chi ha autorità di decidere quali dati includere? Come vengono gestiti gli aggiornamenti? Quale è il ciclo di vita di un dato? Queste domande devono avere risposte chiare e documentate.

A questo punto, creare un *LLM custom* vale davvero la spesa? La risposta dipende dalla situazione specifica, ma per molte realtà di media e grande dimensione, è un sì deciso. Lo sforzo iniziale per sviluppare un *LLM custom* varia a seconda della complessità del dominio, del volume di dati, e del livello di customizzazione desiderato. Tuttavia,

questo sforzo va considerato in prospettiva. Se si utilizza un LLM pubblico inserendo continuamente dati riservati per ottenere performance accettabili, si sta già pagando un prezzo nascosto – la perdita di controllo sulla vostra proprietà intellettuale. Se un modello generico fornisce risposte inaccurate e bisogna dedicare risorse umane a correggere continuamente gli errori, si sta pagando con inefficienza. Dipendere da un *provider* esterno per ogni operazione, rende vulnerabili a cambi di *pricing*, a interruzioni di servizio, a modifiche unilaterali delle politiche. Un *LLM custom*, invece, rappresenta *un asset che cresce di valore nel tempo*. Man mano che viene utilizzato, può essere continuamente migliorato con nuovi dati. Diventa progressivamente più intelligente nel vostro contesto specifico. E rimane, sempre e comunque, vostra proprietà.

Come già detto, per molte realtà, specialmente in una fase iniziale, l'utilizzo pragmatico di modelli pubblici ben scelti rappresenta una strada ragionevole. Tuttavia, nel momento in cui un'organizzazione scopre che l'IA è veramente utile per le sue operazioni, che la utilizzano regolarmente e che la qualità dei risultati è critica per il suo funzionamento, il calcolo cambia. A quel punto, investire in un *LLM custom* – addestrato sui dati certificati, proprietari e puliti dell'organizzazione – non è un lusso bensì una necessità strategica.

L'INTELLIGENZA ARTIFICIALE NEL GOVERNATORATO: GLI INTERVENTI FORMATIVI

La crescita dirompente del fenomeno *Intelligenza Artificiale* è inarrestabile ed incontrovertibile. L'Intelligenza Artificiale rappresenta una vera e propria rivoluzione, uno stravolgimento profondo che sta entrando simultaneamente negli uffici, nelle organizzazioni, nelle scuole e nelle famiglie. È esattamente per questo motivo che il Governatorato dello Stato della Città del Vaticano ha sviluppato una strategia consapevole e responsabile per affrontare questa sfida epocale.

Quando l'Intelligenza Artificiale entra in una *società*, non lo fa in modo uniforme. Ogni contesto – lavorativo, educativo, familiare – vive questa trasformazione in modo distinto, con sfide specifiche e opportunità particolari. Comprendere questa diversità è fondamentale per sviluppare risposte appropriate e consapevoli.

Nel contesto lavorativo, l'introduzione dell'Intelligenza Artificiale richiede attenzione su molteplici fronti. È fondamentale controllare rigorosamente l'autenticità dei contenuti generati dall'IA, proteggendo al contempo la creatività umana che rimane insostituibile. Vanno inoltre affrontati rischi emergenti come quello della *shadow AI*, ovvero l'uso di sistemi di Intelligenza Artificiale non autorizzati o non controllati che operano negli interstizi dei processi lavorativi. L'adozione di attente verifiche sui processi aziendali e protezione dei dati, che oggi rappresentano asset sempre più critici e preziosi per qualsiasi organizzazione moderna, diventa imprescindibile.

Nel contesto scolastico e formativo, la sfida è ancora più profonda. È necessario ripensare completamente il modello didattico tradizionale: dall'offerta formativa alle metodologie di insegnamento che hanno caratterizzato l'educazione fino a oggi. La comunità educativa si interroga legittimamente su come l'Intelligenza Artificiale influenzi la capacità

critica degli studenti. Alcuni esperti temono che la possibilità di delegare attività di ricerca e elaborazione all'IA possa impoverire il pensiero critico dei giovani, essenziale per lo sviluppo di cittadini consapevoli e autonomi.

Nel contesto più intimo e delicato – quello familiare – l'Intelligenza Artificiale tocca aspetti molto più vulnerabili della società. Genitori e adulti si trovano a utilizzare *chatbot* e assistenti virtuali per affrontare questioni che tradizionalmente erano dominio di psicologi, professionisti del settore educativo e specialisti. Questo avviene in un momento in cui l'età media per la prima introduzione di uno *smartphone* nella vita di un bambino si aggira intorno ai dieci anni. La domanda cruciale diventa: come proteggere i minori quando vengono esposti a tecnologie che non comprendono pienamente, e come guidare gli adulti nell'uso responsabile di questi strumenti nel contesto familiare?

Di fronte a questi interrogativi multiformi, una risposta comune emerge chiaramente: è necessario creare consapevolezza specifica attraverso sensibilizzazione e alfa-

betizzazione. In altre parole, la formazione diventa lo strumento principale per trasformare il rapporto tra le persone e la tecnologia.

Il Governatorato ha riconosciuto questa necessità e, da più di un anno, sta investendo significativamente in

programmi formativi specifici sull'Intelligenza Artificiale. Questi programmi non si limitano a spiegare come usare gli strumenti, bensì affrontano temi complessi e articolati. Includono la storia dello sviluppo dell'Intelligenza Artificiale, con i suoi cicli di ottimismo e scetticismo – i cosiddetti “inverni e primavera dell'IA”. Affrontano i limiti intrinseci della tecnologia, le aspettative realistiche, le potenzialità autentiche, e le fallacie a cui l'Intelligenza Artificiale è naturalmente esposta.



Il catalogo formativo è strutturato su più livelli, ognuno pensato per esigenze diverse. Il livello base è rivolto a coloro che stanno muovendo i primi passi nel mondo dell'Intelligenza Artificiale, desiderosi di acquisire familiarità con la terminologia e i contesti applicativi. I livelli più avanzati permettono di entrare in contatto diretto con le varie piattaforme e i diversi modelli di linguaggio (LLM). Attraverso un'esperienza pratica, i partecipanti cominciano a comprendere come identificare lo strumento più adatto a seconda dell'obiettivo specifico che si desidera raggiungere.

Per la natura stessa della materia, i programmi e i contenuti formativi non sono statici. Dato il ritmo accelerato del cambiamento tecnologico, il catalogo formativo è in continuo e costante aggiornamento, rispecchiando sempre la dinamicità della materia stessa. Questo approccio fluido garantisce che la formazione rimanga rilevante e contemporanea.

Il Governatorato non affronta la sfida dell'IA, isolatamente. L'Intelligenza Artificiale è un fenomeno globale che richiede soluzioni coordinate a livello internazionale. Per questo motivo, il Governatorato è in collegamento, attivamente con organizzazioni internazionali capaci di riunire centinaia di paesi nel dialogo su questioni tecnologiche critiche.

Tra queste iniziative, riveste particolare importanza la partecipazione al *World Summit on the Information Society* (WSIS) e alle sue emanazioni recenti. In questo contesto, il Governatorato è coinvolto in iniziative come *AI for Good*, un evento e una piattaforma il cui obiettivo è identificare tecnologie, metodologie e strategie finalizzate all'adozione consapevole e responsabile dell'Intelligenza Artificiale in qualsiasi contesto applicativo.

Queste strategie globali necessariamente affrontano problemi già noti e ben documentati. Primo fra tutti, il divario digitale (*digital divide*), ovvero l'accesso diseguale alla tecnologia tra paesi ricchi e paesi in via di sviluppo. Inoltre, si affronta con rigore il tema dei pregiudizi algoritmici (*bias*), e soprattutto il pericolo concreto della discriminazione che potrebbe derivare dall'assenza di modelli di linguaggio tradotti in tutte le lingue mondiali. Se l'Intelligenza Artificiale è addestrata prevalentemente in inglese e in poche altre lingue europee, gli utenti di lingue meno rappresentate potrebbero subire prestazioni significativamente inferiori, perpetuando e amplificando le ineguaglianze globali già esistenti.

Oltre alla formazione e alla partecipazione internazionale, il Governatorato continua ad investire in un'altra area cruciale: la sovranità e la gestione responsabile dei dati. Questo tema assume importanza particolare se si considerano i dati come il "petrolio" del ventunesimo secolo – una risorsa straordinariamente preziosa da creare, mantenere, difendere e valorizzare.

Per un'entità come il Governatorato, il controllo e la protezione della sovranità dei dati non è una questione secondaria, bensì una caratteristica imprescindibile. Non è accettabile che dati sensibili riguardanti le operazioni dello Stato vengano gestiti da server esterni controllati da aziende private straniere. È necessario sviluppare capacità interne di gestione e protezione.

La soluzione che il Governatorato intende valutare ed adottare, sposa questo approccio: sviluppare soluzioni di Intelligenza Artificiale locali e interne. Questo significa dotarsi di *hardware* adeguato in termini di potenza computazionale, ma anche attento ai risvolti energetici e alla sostenibilità ambientale. L'obiettivo è elaborare i dati interni in modo autonomo, valorizzarli e renderli fruibili secondo i nuovi paradigmi della *generative AI*, senza dipendere da infrastrutture esterne.

Questo approccio presenta sfide significative. Si tratta di scenari nuovi, in gran parte inesplorati, ricchi di insidie tecniche e organizzative. Tuttavia, allo stesso tempo, contengono grandi opportunità. Rappresentano un modello alternativo rispetto alla centralizzazione del potere tecnologico nelle mani di poche realtà globali. Dimostrano che è possibile sviluppare Intelligenza Artificiale responsabile, eticamente consapevole e sovrana anche per realtà di dimensioni più contenute.

La tecnologia non è neutra, le scelte fatte oggi determineranno il panorama del domani e la responsabilità nel governare l'innovazione ricade su tutti gli attori della società, dalle istituzioni ai singoli individui.

Il Governatorato sta lavorando attivamente per raggiungere una configurazione stabile, contemporanea e pronta ad affrontare pienamente l'era dell'Intelligenza Artificiale. Questo non è compito che possa essere delegato a una sola parte dell'organizzazione. Richiede il contributo e la collaborazione consapevole di tutti. Richiede che ognuno sviluppi una comprensione adeguata dei rischi e delle opportunità, e che agisca con integrità e responsabilità etica nelle proprie scelte quotidiane.

La visione è che sia assolutamente possibile abbracciare l'Intelligenza Artificiale senza sacrificare i valori fondamentali di sovranità, etica, equità e trasparenza. Formazione consapevole, partecipazione globale e sovranità tecnologica quali fondamenta su cui costruire innovazione.

D. G.

LE COMMISSIONI DEL GOVERNATORATO: STRUMENTI DI GOVERNANCE SPECIALIZZATA

Le Commissioni del Governatorato nascono per coadiuvare gli Organi di Governo in settori che richiedono expertise particolare, garantendo al contempo trasparenza, collegialità nelle decisioni e conformità ai principi etici che caratterizzano l'operato dello Stato. Nel corso dei decenni sono state istituite commissioni dedicate a materie quali il personale, la disciplina, le questioni monetarie e la selezione dei collaboratori laici, ciascuna con funzioni consultive, deliberative o di vigilanza a seconda dell'ambito di competenza.

In questo contesto di *governance* specializzata si inserisce la recente istituzione della Commissione sull'Intelligenza Artificiale, prevista dal Decreto n. DCCII del 30 dicembre 2024 ed entrata in vigore il 1° gennaio 2025. Di fronte alla rapida diffusione dei sistemi di Intelligenza Artificiale e alla loro crescente influenza su ogni aspetto della vita sociale e istituzionale, il Governatorato ha ritenuto necessario dotarsi di un organismo dedicato che garantisca un utilizzo etico, trasparente e responsabile di questa tecnologia.

La Commissione ha il compito di predisporre leggi e regolamenti attuativi delle linee guida sull'Intelligenza Artificiale, esprimere pareri sulle proposte di sperimentazione e applicazione di sistemi di IA, svolgere attività di monitoraggio continuo segnalando potenziali rischi, e redigere relazioni

semestrali sull'impatto dell'uso dell'Intelligenza Artificiale nello Stato vaticano. In sostanza, la Commissione funge da bussola orientativa per garantire che lo sviluppo tecnologico rimanga sempre al servizio della dignità umana e del bene comune, in piena coerenza con i valori fondamentali della Santa Sede.

La Commissione è composta da cinque membri nominati dal Presidente del Governatorato e presieduta dal Segretario Generale, come definito dall'art.14 del Decreto. I membri provengono da tre direzioni strategiche: l'Ufficio Giuridico, la Direzione delle Telecomunicazioni e dei Sistemi Informatici, e la Direzione dei Servizi di Sicurezza e Protezione Civile. Questa composizione multidisciplinare assicura che le decisioni tengano conto degli aspetti legali, tecnologici e di sicurezza. Il mandato ha durata triennale rinnovabile.

La Commissione si è già riunita ed è attivamente al lavoro per mettere in pratica i primi passi operativi. La tempestività dell'azione riflette la consapevolezza della delicatezza della materia e dell'importanza di disporre di un assetto concreto sin da subito, capace di rispondere efficacemente alle sfide che l'Intelligenza Artificiale pone quotidianamente all'amministrazione vaticana.

D. G.



