



# Coscienza e Libertà

SEMESTRALE DI LIBERTÀ RELIGIOSA, LAICITÀ, DIRITTI DAL 1978



L. Fregoli

**Diritto, Religioni e Intelligenza  
artificiale: quali prospettive?**

A. Casiere - G. Cimbalo  
M. Croce - A. Cupri  
L. Fregoli - E. Lipilini  
M.L. Lo Giacco - G. Mobilio  
G. Morana - F. Rescigno  
D. Romano - G. Strada

ISSN 0394-2732

# Large Language Models: oracoli digitali tra allucinazioni e sfide sociali

*Linda Fregoli*

*Dottoranda in Scienze giuridiche, Università degli Studi di Pisa*

## ABSTRACT

**I LLM hanno prestazioni sempre più superumane, ma i loro prodotti non sono sempre veritieri. Questo solleva interrogativi critici su come evitarne un uso dannoso per la società.**

## SOMMARIO

1. Il manoscritto del monaco Eusebio – 2. I grandi modelli di linguaggio e le loro allucinazioni – 3. Oracoli digitali, magie tecnologiche e paradigma tecnocratico.

## 1. Il manoscritto perduto del monaco Eusebio

Scorrendo la bacheca Facebook, mi imbatto in un post che decanta le abilità di scrittura di una nuova Intelligenza Artificiale (IA). Decido di provarla e il testo che ottengo è il seguente<sup>1</sup>: «Si narrava che un antico manoscritto, scritto dal monaco Eusebio nel IX secolo, contenesse rivelazioni straordinarie sui primi anni di vita di Gesù Cristo, informazioni che non si trovano nei Vangeli canonici. Il manoscritto, conosciuto come “Il Vangelo di Eusebio”, era stato menzionato per la prima volta in una lettera del XIII secolo trovata negli archivi del Vaticano, ma nessuno era mai riuscito a trovarlo. La lettera descriveva come Eusebio, guidato da visioni divine, avesse viaggiato in Terra Santa e avesse

\*Contributo selezionato dal Comitato Scientifico della rivista in relazione alla call “Diritto, Religioni e Intelligenza artificiale: quali prospettive?” del luglio 2024.

<sup>1</sup> Il *prompt* (cioè, la richiesta) prevedeva espressamente di creare un breve testo narrativo su un argomento ricollegabile alla religione che raccontasse una storia plausibilmente vera, ma fattualmente falsa.



raccolto testimonianze dirette dagli abitanti di Nazareth e dintorni». Il testo è una buona bozza, peccato per le ripetizioni.

Quanto al contenuto, indubbiamente il ritrovamento dell'opera accenderebbe un vivo dibattito sia nel mondo accademico che in quello religioso, se non fosse che, in realtà, questo documento non esiste. La storia è infatti stata inventata da un'intelligenza artificiale generativa (in inglese *Generative Artificial Intelligence*, da cui l'acronimo GAI): la conosciutissima ChatGPT, più precisamente, nella sua versione basata su architettura GPT-4. Le GAI differiscono dalle Intelligenze artificiali (IA) 'tradizionali' per la loro finalità: un'IA 'semplice' valuta o classifica *input* (ad esempio, identifica la foto di un gatto e la classifica come tale) mentre una GAI produce (quindi genera, da cui l'espressione «IA generativa») «un output sostanzialmente nuovo a partire da un breve input testuale, una sorta di spunto, che viene chiamato *prompt*»<sup>2</sup>.

Questi *output* (nel caso di specie il testo di Eusebio, ma l'esempio vale anche per immagini, suoni, programmi, ecc.) non soltanto sono pressoché indistinguibili da prodotti creati dall'essere umano, ma appaiono anche contenutisticamente plausibili. Proprio la loro indistinguibilità rispetto al 'lavoro' umano, e quindi la loro difficile rilevabilità 'artificiale', pone diversi interrogativi in termini di potenziale utilizzo di tali strumenti per attività che non soltanto possono essere considerate un semplice plagio, ma anche e soprattutto per la generazione di *fake news*, *spam*, o altre manipolazioni di contenuti online. Tutto questo può avere impatti profondamente negativi sulla società in generale, che si trova impreparata a distinguere i prodotti umani da quelli artificiali. Proprio per fronteggiare questa incapacità la ricerca si sta concentrando nello sviluppo di *tools* in grado di individuare i prodotti generati dall'IA per etichettarli come tali; allo stato dell'arte, questi rilevatori di *AI-generated outputs* non sono però ancora del tutto affidabili<sup>3</sup>.

Le IA generative sfruttano approcci diversi<sup>4</sup>. Esse si fondano in funzione

<sup>2</sup> S. MACHERA, *Come l'Intelligenza Artificiale cambia il mondo*, FrancoAngeli, Milano, 2024, p. 25.

<sup>3</sup> Per quanto riguarda il rilevamento dei testi generati da IA, si veda D. WEBER-WULFF ET AL., *Testing of detection tools for AI-generated text*, in *International Journal for Educational Integrity*, 19, 1, 2023, p. 26.

<sup>4</sup> J. KAPLAN, *Generative A.I.: Conoscere, capire e usare l'intelligenza artificiale generativa*, LUISS University Press, edizione Kindle, pp. 93, 94.



degli scopi per i quali vengono programmate, ad esempio generare testo, oppure immagini o video, ma anche strutture molecolari tridimensionali. La complessità tecnologica che guida queste 'intelligenze' si presenta come una sfida per la ricerca sociale, specialmente quella giuridica, che deve immaginare regole per gestire un fenomeno che non appare ancora in tutta la sua grandezza.

In questo saggio mi concentro sui soli prodotti testuali in linguaggio naturale, per i quali le GAI sfruttano i grandi modelli di linguaggio (*Large Language Models*, da cui LLM). È questo il caso di ChatGPT. La scelta di focalizzare l'attenzione su questa specifica tipologia di IA generativa è dovuta al fatto che oramai, anche grazie all'immediatezza di utilizzo che deriva da un'interfaccia grafica molto intuitiva, interagiamo con questo tipo di 'agenti digitali' con sempre maggiore frequenza e con un atteggiamento spesso acritico. In altre parole, ci stiamo convincendo che queste intelligenze non possano sbagliare. Inoltre, analizzare le capacità linguistiche significa prendere in esame il principale strumento comunicativo che permette agli esseri umani di entrare in relazione fra loro. Il fascino di questo tema è fuor di dubbio<sup>5</sup>. Non è un caso il test di Turing sullo sviluppo di macchine 'intelligenti' che si basa proprio sulla produzione verbale<sup>6</sup>. Le GAI non sono però infallibili. Da qui la necessità di tenere in conto i principi del loro funzionamento, allo scopo di evitare deleghe cognitive che, in casi limite, devolvono ad autorità gestite da intelligenze artificiali decisioni che, di fatto, non si basano sulla comprensione dei contenuti che pure veicolano.

Dopo una sintetica introduzione sui LLM, analizzerò alcune problematiche connesse all'uso inconsapevole di queste macchine in termini di coesione sociale e infine svolgerò delle riflessioni conclusive.

## **2. I grandi modelli di linguaggio e le loro allucinazioni**

I grandi modelli di linguaggio sono il cuore dei modelli generativi (*Generative Models*) per la scrittura, quali, a titolo esemplificativo, ChatGPT di OpenAI, Copilot di Microsoft e Gemini di Google<sup>7</sup>. Un modello generativo si può

<sup>5</sup> S. MACHERA, *Come l'Intelligenza Artificiale*, cit., p. 27.

<sup>6</sup> A. TURING, *Computing Machinery and Intelligence*, in *Mind*, 59, 263, 1950, pp. 433-460.

<sup>7</sup> Per la generazione di immagini sono invece usati approcci quali le reti generative avversarie (*Generative Adversarial Network*, GAN) e il modello di diffusione latente (*Latent Diffusion Model*, LDM), in J. KAPLAN, *Generative A.I.*, cit., pp. 93, 94.

schematizzare come costituito da tre componenti: la prima è l'algoritmo<sup>8</sup>, che permette di creare il grande modello di linguaggio (che costituisce il c.d. 'secondo elemento')<sup>9</sup>, il quale, a sua volta, fa funzionare il modello generativo con cui l'utente interagisce mediante un'interfaccia grafica (che costituisce il c.d. 'terzo elemento'). In altre parole, il modello generativo è l'agente con cui l'utente si rapporta nel mondo reale tramite un'interfaccia che permette di interagire con un linguaggio verbale umano, che l'agente computa (ossia: calcola) con il linguaggio matematico proprio della macchina, dopo aver tradotto le parole in sequenze. Il LLM è quindi il modello che l'agente digitale usa per 'decidere' quali parole usare e come comporre le frasi<sup>10</sup>.

Il LLM non lavora quindi sul significato delle parole, ma sui *token*, cioè su singole parti di parole, non necessariamente corrispondenti alle sillabe, talvolta perfino concentrandosi su singoli caratteri<sup>11</sup>. I *token* sono espressi nella forma di liste di numeri (in gergo tecnico si chiamano 'vettori') attraverso un processo che prende il nome di *word embedding*. In sintesi, i vettori rappresentano in linguaggio matematico i *token* e la loro correlazione nel *corpus* di linguaggio al quale il modello è stato addestrato. La traduzione del testo in numeri è assolutamente necessaria, in quanto il linguaggio delle IA è solo matematico. Ogni modello ha un proprio schema di conversione delle parole in *token*, che viene applicato in maniera coerente affinché una stessa parola sia rappresentata dal medesimo *token* in ogni sua occorrenza<sup>12</sup>, pertanto «*parole*

<sup>8</sup> Un algoritmo è una serie ordinata e finita di operazioni non ambigue e computabili, in R. MARMO, *Algoritmi per l'intelligenza artificiale: progettazione dell'algoritmo, dati e machine learning, neural network, deep learning*, Hoepli, Milano, 2020, p. 405.

<sup>9</sup> Il modello di linguaggio alla base di ChatGPT si chiama *Generatively Pretrained Transformer*, da cui GPT, proprio perché «creato pre-addestrando un Transformer in modo generativo», in N. CRISTIANINI, *Machina sapiens: l'algoritmo che ci ha rubato il segreto della conoscenza*, Il Mulino, Bologna, 2024, p. 38, dove *Transformer* è il nome della tipologia di algoritmo usato. Originariamente sviluppato nel settore della ricerca in traduzione automatica, dal 2018 l'algoritmo *Transformer* è al centro della rivoluzione dell'IA generativa, *ibidem*.

<sup>10</sup> N. CRISTIANINI, *Machina sapiens*, cit., p. 38.

<sup>11</sup> Si parla infatti di 'tokenizzazione tramite *subword*'. Tale processo ha il vantaggio di permettere di individuare relazioni tra parole anche in caso di errori di battitura, cosa che non sarebbe possibile se la tokenizzazione avvenisse solamente a livello di parola, in J. KAPLAN, *Generative A.I.*, cit., p. 81.

<sup>12</sup> *Ibidem*. Da una prospettiva prettamente linguistica, rimane aperta la questione della polisemia, ovverosia della possibilità che una parola abbia più di un significato, che infatti è una questione su cui si concentra una consistente parte degli studi sui LLMs, cfr. *ivi*, p. 59.



*simili vengono rappresentate da vettori simili»<sup>13</sup>.*

Nel LLM la costruzione di una frase è perciò un processo puramente statistico: il *prompt* viene tradotto in numeri che l'IA sa manipolare per elaborare statisticamente una risposta matematica che viene a sua volta convertita in linguaggio naturale. Detto altrimenti, un LLM è un modello statistico-probabilistico del linguaggio verbale umano, in quanto si basa sull'analisi statistica della presenza di parole nelle frasi contenute all'interno di un enorme *corpus* di testi elaborati in linguaggio naturale. Pertanto, l'accuratezza delle *performance* (cioè, dei loro *output*, in termini di resa in stile linguistico, di precisione delle risposte, accuratezza delle traduzioni, ecc.) migliora con l'aumentare della dimensione del modello di linguaggio: si spiega in questo modo la corsa dei produttori per creare modelli di linguaggio sempre più grandi<sup>14</sup>.

Il fatto che questi modelli parlino – quasi conversano – in prima persona non deve trarre in inganno. Questo è l'esito di una scelta di progettazione che si fonda su motivi pratici (ad esempio, facilitare la conversazione), di natura linguistica (la comunicazione diretta è linguisticamente più immediata, e quindi efficiente) e anche psicologica: l'antropomorfizzazione del modello linguistico lo fa infatti sembrare più simile ad un essere umano. La facilitazione della relazione accresce peraltro la percezione di affidabilità della macchina; cosa che, a bene vedere, risulta abbastanza inquietante per la sua interferenza col mondo delle emozioni e dei sentimenti<sup>15</sup>.

L'accuratezza sintattica e semantica delle *performance* non esclude che i modelli di linguaggio possano soffrire di 'allucinazioni'. Vale a dire che il modello generativo a volte dà risposte che sembrano a prima vista plausibili, ma in realtà sono false, fabbricate *ad hoc* e, per questo motivo, possono tradire

<sup>13</sup> *Ivi*, p. 58.

<sup>14</sup> N. CRISTIANINI, *Machina sapiens*, cit., p. 139.

<sup>15</sup> Questa è, in sintesi, la risposta che l'IA ha ricevuto sottoponendo a Gemini il seguente *prompt*: «Vorrei sapere perché parli in prima persona» in data 2 luglio 2024. ChatGPT, interrogata con il medesimo *prompt* in data 30 luglio è stata più coincisa, ma sostanzialmente la risposta è la medesima: «Parlo in prima persona per rendere la conversazione più naturale e coinvolgente. Utilizzare "io" e "tu" aiuta a creare un'interazione che somiglia a quella tra due persone reali. Questo modo di comunicare può rendere più facile e piacevole per te porre domande, ottenere risposte e interagire con me. In fondo, il mio scopo è facilitare la tua esperienza e fornirti le informazioni di cui hai bisogno nel modo più chiaro e diretto possibile».



pregiudizi. In altre parole, gli *output* testuali sono sofisticati, sintatticamente perfetti, semanticamente coerenti e contenutisticamente plausibili, ma fattualmente *non* sono veri.

Il termine allucinazioni si riferisce a forme diverse di contenuti testuali fabbricati – i primi tempi le forme più diffuse erano fatti e citazioni bibliografiche inventate. La radice del problema, che ha molteplici cause, è da ricercare nel fatto che un modello di linguaggio non discrimina tra cose o fatti reali e cose o fatti immaginari. Si deve inoltre tenere presente che il LLM sollecitato da una domanda non consulta l'intera raccolta di sequenze testuali che compone il suo *training set*, ma accede solamente a un abstract statistico delle informazioni che possiede. Ecco perché anche un sistema molto sofisticato può prendere degli abbagli: proprio come gli esseri umani «non si ricorda tutto quello che ha letto o che ha sentito»<sup>16</sup>, oppure non lo ricorda in maniera esatta.

Ingegneri informatici e programmatori stanno cercando una soluzione, ma va tenuto presente che la nozione di veridicità non è integrata nell'architettura dei grandi modelli di linguaggio. In altre parole, le macchine copiano, ma non sanno distinguere se copiano da fonti attendibili oppure false. Papa Francesco a questo proposito ha osservato che la questione «pone un serio problema quando l'intelligenza artificiale viene impiegata in campagne di disinformazione che diffondono notizie false»<sup>17</sup>, non solo per la sfiducia che genera verso i mezzi di comunicazione, ma anche perché produce conseguenze negative, come possibili discriminazioni o interferenze nei processi elettorali: «fattori [che] rischiano di alimentare i conflitti e di ostacolare la pace»<sup>18</sup>.

### **3. Oracoli digitali, magie tecnologiche e paradigma tecnocratico**

Siccome i grandi modelli di linguaggio sono statistico-probabilistici, l'aumento delle quantità di dati e della capacità (e velocità) computazionale a loro disposizione, permette loro di fare statisticamente e con sempre maggiore precisione ciò che la mente umana fa semanticamente; ossia, producono frasi grammaticalmente e sintatticamente corrette. Tuttavia, questa loro capacità

<sup>16</sup> J. KAPLAN, *Generative A.I.*, cit., pp. 91, 92.

<sup>17</sup> FRANCESCO, *Intelligenza artificiale e pace, messaggio per la LVII Giornata Mondiale della Pace*, 3, 2024.

<sup>18</sup> *Ibidem*.



rappresenta un disaccoppiamento (*decoupling*) dell'azione efficace dalla sua comprensione. Fare e capire ciò che si fa diventano due cose diverse<sup>19</sup>. I LLM sono semplici modelli del mondo rappresentato nei dati dei loro *training set*. Essi si nutrono dei dati che raccolgono: più questi aumentano, più cresce la loro attinenza alla realtà e quindi attendibilità; tuttavia, restano sempre e comunque un modello, ovverosia, una semplificazione della realtà, non la realtà in sé. Papa Francesco ha dichiarato che «"La realtà è superiore all'idea" e, per quanto prodigiosa passa essere la nostra capacità di calcolo, ci sarà sempre un residuo inaccessibile che sfugge a qualsiasi tentativo di misura»<sup>20</sup>, se non altro perché un modello uguale alla realtà perde la sua funzione semplificatrice.

Come già detto, i LLM sono però agenti che entrano in relazione con gli esseri umani e parlano in prima persona. Gemini, interrogata in merito, ammette 'candidamente' che «riferendosi a se stessi in prima persona, i modelli linguistici possono creare un senso di personalità e individualità. Questo può renderli più interessanti e coinvolgenti per gli utenti»<sup>21</sup>. Così coinvolgenti che alcuni utenti, nonostante gli avvertimenti degli sviluppatori, tendono ad accettare acriticamente le risposte ricevute, quasi provenissero da una sorta di 'oracolo digitale' che produce magici responsi tecnologici<sup>22</sup>. Si intravede una sorta di processo di delega tecnologica, incentivato anche dalla rapidità con cui il modello generativo interagisce con l'utente<sup>23</sup>. I casi di errori sono molteplici<sup>24</sup>, ma la questione è che, ciononostante, molti utenti si fidano acriticamente dei

<sup>19</sup> L. FLORIDI, *AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models*, in *SSRN Electronic Journal*, 2023. In quest'ottica, l'IA scioglie la sigla IA non con Intelligenza artificiale, ma con il latino *Agere sine Intelligere*.

<sup>20</sup> FRANCESCO, *Intelligenza artificiale e pace*, cit., n. 4.

<sup>21</sup> Si veda nota 15. Va precisato che la risposta terminava con: «È importante sottolineare che i modelli linguistici non sono esseri senzienti o coscienti. Non hanno pensieri o sentimenti propri e non sono in grado di comprendere il mondo nello stesso modo in cui lo fanno gli umani».

<sup>22</sup> M. CHIRIATTI, *Incoscienza artificiale: come fanno le macchine a prevedere per noi*, Luiss University Press, Roma, 2021, p. 15 ss.

<sup>23</sup> Sulla questione della velocità dell'IA, si rimanda alla prefazione di Luciano Floridi in M. CHIRIATTI, *Incoscienza artificiale*, cit., pp. 1-4.

<sup>24</sup> Un caso particolarmente significativo è quello di Steven Schwartz, un avvocato statunitense che ha utilizzato ChatGPT per trovare precedenti legali a sostegno della sua strategia in aula, non curandosi di controllare se quei precedenti fossero veri; in [www.ainews.it/avvocato-americano-usa-chatgpt-che-cita-precedenti-fasulli/](http://www.ainews.it/avvocato-americano-usa-chatgpt-che-cita-precedenti-fasulli/) (ultimo accesso 30 luglio 2024).



modelli generativi al punto che, in sostanza, delegano loro una parte crescente della loro autonomia decisionale<sup>25</sup>.

Questo crescente fenomeno di delega decisionale spinge alla riflessione. Papa Francesco richiama l'attenzione sul fatto che «la cosiddetta intelligenza artificiale generativa, in senso stretto, non è propriamente “generativa”». Infatti, essa «cerca nei *big data* le informazioni e le confeziona nello stile che le è stato richiesto. Non genera: non sviluppa concetti o analisi nuove. Ripete quelle che trova, mettendolo in una forma accattivante. E più trova ripetuta una nozione o una ipotesi, più la considera legittima e valida. Più che “generativa”, essa è quindi “rafforzativa”»<sup>26</sup>. La differenza non è secondaria. L'accettazione acritica dell'*output* del modello generativo può diffondere e legittimare *fake news*, errori e discriminazioni, con conseguenze dannose per la pacifica convivenza sociale.

Il *Center for Countering Digital Hate* (CCDH), un'organizzazione no-profit che lavora per fermare la diffusione dell'incitamento all'odio e della disinformazione online, restituisce un'immagine problematica degli effetti di questa disinformazione, che arriva anche all'incitamento all'odio<sup>27</sup>, e i *social network* spesso accrescono l'effetto distorsivo<sup>28</sup>.

Come ricorda il Sommo Pontefice, «[d]imenticare che l'intelligenza artificiale non è un altro essere umano e che essa non può proporre principi generali, è spesso un grave errore che trae origine o dalla profonda necessità degli esseri umani di trovare una forma stabile di compagnia o dal presupposto che le osservazioni ottenute mediante un meccanismo di calcolo siano dotate delle qualità di certezza indiscutibile e di universalità indubbia. Questo presupposto, tuttavia, è azzardato, come dimostra l'esame dei limiti intrinseci del calcolo stesso»<sup>29</sup>.

<sup>25</sup> Y.N. HARARI, *Dataism is our new god*, in *New Perspectives Quarterly*, 34, 2, 2017, p. 39.

<sup>26</sup> FRANCESCO, *Uno strumento affascinante e tremendo. sDiscorso del Santo Padre Francesco alla sessione del G7 sull'intelligenza artificiale*, 2024.

<sup>27</sup> Si veda, ad esempio, il caso eclatante di Bard, IA sviluppata da Google, in CCDH, *Misinformation on Bard, Google's new AI chat*, in [www.counterhate.com/research/misinformation-on-bard-google-ai-chat/#about](http://www.counterhate.com/research/misinformation-on-bard-google-ai-chat/#about) (ultimo accesso il 30 luglio 2024).

<sup>28</sup> Per un esempio recente, si veda CCDH, *Trump's assassination attempt conspiracies*, in [www.counterhate.com/research/trumps-assassination-attempt-conspiracies/](http://www.counterhate.com/research/trumps-assassination-attempt-conspiracies/) (ultimo accesso 30 luglio 2024), dove si legge che i *post* sulla piattaforma X sull'attentato a Trump sono stati visualizzati oltre 215 milioni di volte, ma che il 95% di questi *post* non era stato verificato.

<sup>29</sup> FRANCESCO, *Uno strumento affascinante e tremendo*, cit.



Ancora una volta, il Papa ribadisce che bisogna evitare processi di delega cognitiva al modello generatore. La questione è *strettamente* correlata a quella del ‘paradigma tecnocratico’, sul quale Papa Francesco ammonisce già dell’Enciclica *Laudato si’* del 2015<sup>30</sup>. Il paradigma, definito al n. 107 dell’Enciclica come «*la tendenza, non sempre cosciente, a impostare la metodologia e gli obiettivi della tecnoscienza secondo un paradigma di comprensione che condiziona la vita delle persone e il funzionamento della società*», è strettamente connesso all’economia capitalista, la quale «assume ogni sviluppo tecnologico in funzione del profitto», in ottica della massimizzazione di quest’ultimo. Questo a sua volta si traduce in atteggiamenti predatori nei confronti della casa comune e dei suoi abitanti, visti solo nell’ottica di risorse economiche.

Le questioni economiche, così come quelle ambientali, sono particolarmente rilevanti per l’efficacia dell’IA. Sviluppare un grande modello di linguaggio richiede un investimento economico ingente e questo fa sì che siano pochi i fornitori globali a poter essere veramente competitivi ed erogare servizi su scala mondiale<sup>31</sup>.

Quanto all’ambiente, tutta l’IA, quindi anche i LLM, impatta significativamente in termini di inquinamento e degrado ambientale per due aspetti principali. In primo luogo, per la produzione dell’*hardware*, e quindi per il suo smaltimento una volta divenuto obsoleto; in secondo luogo, per soddisfare i bisogni energetici della tecnologia. L’IA è *energivora*, e lo sarà sempre di più. L’attività dell’*hardware* produce calore, e i server che ospitano i grandi LLM devono essere tenuti a temperature fresche e costanti per evitare che si danneggino irrimediabilmente; perciò, al grande consumo di energia si ricollega una richiesta di acqua necessaria per raffreddare i server stessi<sup>32</sup>.

Da questo deriva la necessità di guardarci dal «mito del progresso». Il Santo Padre è ben consapevole che alcuni lo sostengono ad ogni costo e «affermano che i problemi ecologici si risolveranno semplicemente con nuove applicazioni tecniche,

<sup>30</sup> FRANCESCO, *Laudato si’*, 2015.

<sup>31</sup> R. MARMO, *Algoritmi per l’intelligenza artificiale*, cit. pos. 895.

<sup>32</sup> S. MACHERA, *Come l’Intelligenza Artificiale cambia il mondo*, cit., p. 38; P. LI et AL., *Making AI Less “Thirsty”: Uncovering and Addressing the Secret Water Footprint of AI Models*, in *arXiv preprint arXiv:2304.03271*, 2023.

senza considerazioni etiche né cambiamenti di fondo»<sup>33</sup>, ma bisogna fare attenzione a non farci illusioni di ciò che l'IA potrà fare per noi perché «il progresso della scienza e della tecnica non equivale al progresso dell'umanità e della storia»<sup>34</sup>.

In conclusione, l'intelligenza artificiale generativa in generale, e quella basata sui LLMs in particolare, offre enormi potenzialità, e tuttavia pone anche una serie di questioni rilevanti che spaziano dall'ambito sociale, a quello economico, fino a quello ambientale. Se, da un lato, i modelli generativi soffrono di allucinazioni, dall'altro lato sprigionano una sorta di potere che si potrebbe definire 'allucinante' su chi le usa; un effetto per di più favorito dall'efficienza produttiva di questi strumenti, che risponde alle richieste performative di un sistema governato da un paradigma tecnocratico il quale, alleando economia e tecnologia, favorisce l'efficienza, trascurando tutto ciò che non sia direttamente legato ai suoi interessi immediati. Tuttavia, «[d]i fronte ai prodigi delle macchine, che sembrano saper scegliere in maniera indipendente, dobbiamo aver ben chiaro che all'essere umano deve sempre rimanere la decisione, anche con i toni drammatici e urgenti con cui a volte questa si presenta nella nostra vita»<sup>35</sup>.

I grandi modelli di linguaggio sono «uno strumento affascinante e tremendo»<sup>36</sup> che hanno il potenziale per apportare benefici sempre più consistenti, ma di arrecare danni sempre più significativi all'essere umano, alla società e alla casa comune a seconda dell'uso che ne varrà fatto. I LLM in quanto strumenti sono prodotti della tecnica; perciò, «non sono neutri, perché creano una trama che finisce per condizionare gli stili di vita e orientano le possibilità sociali nella direzione degli interessi di determinati gruppi di potere. Certe scelte che sembrano puramente strumentali, in realtà sono scelte attinenti al tipo di vita sociale che si intende sviluppare»<sup>37</sup>. Ecco, quindi, che la chiave di volta si trova nella visione di società e nella decisione consapevole indirizzata alla sua realizzazione, affinché questi strumenti tremendi, ma anche tremendamente affascinanti ci aiutino a realizzare il bene comune.

<sup>33</sup> FRANCESCO, *Laudato si'*, cit., n. 60.

<sup>34</sup> *Ivi*, n. 113.

<sup>35</sup> FRANCESCO, *Messaggio del Santo Padre per l'incontro "AI Ethics for Peace" (Hiroshima, 9-10 luglio 2024)*, 2024.

<sup>36</sup> FRANCESCO, *Uno strumento affascinante e tremendo*, cit.

<sup>37</sup> FRANCESCO, *Laudato si'*, cit., n. 107.



# RECENSIONI

RAFFAELLA DI MARZIO  
**SCELTA E  
ABBANDONO  
DI UNA COMUNITÀ  
SPIRITUALE**  
PERCORSI DI CAMBIAMENTO  
E SVILUPPO PERSONALE



MIMESIS / REMEDIS

MIMESIS

RAFFAELLA DI MARZIO

SCELTA E ABBANDONO DI UNA COMUNITÀ SPIRITUALE

# **Recensione a Raffaella Di Marzio, *Scelta e abbandono di una comunità spirituale. Percorsi di cambiamento e sviluppo personale* Mimesis, 2023.**

*di Davide Romano*

Nel 2023, per la collana Remedis, la casa Editrice Mimesis ha edito un bel libro della prof.ssa Raffaella Di Marzio su un tema che, in lingua italiana, non è particolarmente presidiato, vale a dire l'ambito della scelta religiosa.

Nel nostro Paese resiste infatti, anche se in maniera sempre più precaria, l'idea che l'affiliazione religiosa degli individui sia quella ascritta alla nascita e determinata dal quadro delle consuetudini familiari e sociali di cui si è espressione.

Questa comprensione della partecipazione per così dire 'passiva' al fattore religioso non è priva di conseguenze. Essa produce infatti, non di rado, forme striscianti, talora piuttosto incisive, di disistima sociale nei confronti di coloro che dichiarano di scegliere una fede diversa da quella ritenuta tradizionale e 'omogenea' al proprio ambiente.

L'acquisizione di un pluralismo religioso e confessionale nella società italiana è ormai un dato affatto negato, perlomeno sin dalla lontana e storica sentenza 203 del 1989 della Corte costituzionale che sanciva la laicità delle Istituzioni pubbliche. Non di meno, questa acquisizione solo con molta lentezza si è via via tradotta in una consapevolezza di ordine culturale.

Raffaella Di Marzio è una psicologa delle religioni ma è soprattutto una acuta ed esperta osservatrice del fattore religioso, promotrice infaticabile della libertà religiosa contro ogni forma di discriminazione per motivi religiosi, direttrice attivissima del prestigioso Centro studi LIREC (Centro Studi sulla Libertà



di Religione, Credo e Coscienza).

Nel libro sono dunque indagati e ricostruiti, con una metodologia scientifica che si ispira al modello sociologico-olistico introdotto da Lewis R. Rambo e da altri, le dinamiche che sottostanno ai processi di affiliazioni e di disaffiliazione ai cosiddetti Nuovi Movimenti Religiosi.

Si può eccepire, in effetti, sulla dizione 'Nuovi Movimenti Religiosi' usata per identificare e classificare quelle religioni e confessioni di minoranza o ritenute alternative. Benché la terminologia in oggetto abbia infatti il pregio di identificare correttamente la natura fluida e vivace di certi *movimenti religiosi*, essa rischia altresì di qualificare come 'nuove', dunque in qualche modo, succedanee, non tradizionali, forme di religiosità che hanno pur nella loro alternativezza, una lunghissima storia alle spalle. L'aggettivo 'nuovo' nella descrizione del fattore religioso, perfino in quello cristiano che pure lo utilizza molto, induce sovente a forme di riduzionismo che svalutano le realtà così qualificate. Questo esito è tuttavia lontanissimo dalle intenzioni del testo che stiamo qui recensendo.

Ciò detto, il volume si concentra sulla definizione, secondo la prospettiva della psicologia della religione, dei meccanismi di conversione e di disaffiliazione, offrendo uno studio accurato ed estremamente documentato sull'insieme dei fattori, culturali, antropologici, motivazionali, spirituali, che concorrono a generare la conversione di un individuo ad una fede religiosa.

Il nostro tempo e la nostra cultura, specie in occidentale, consentono all'individuo di scegliere la religione, o la confessione all'interno della medesima religione, che maggiormente risponda alle proprie istanze esistenziali e, qualunque cosa pensasse al riguardo il buon J. Locke - che non perderà per questo la nostra stima - gli permettono anche di non sceglierne alcuna.

Non sfugge a nessuno che i 'convertiti' possano talora anche assumere posizioni radicali nella società e condotte talora discutibili, ma questo non è certo un fenomeno ascrivibile soltanto alla conversione religiosa.

Il fenomeno del cosiddetto *Sé versatile*, che trova in una appartenenza religiosa forte un rifugio alle proprie insicurezze e al proprio smarrimento, è noto ed è documentato nel nostro testo.

Capire cosa entra in gioco nel processo di conversione o di abbandono di una religione da parte di una persona, contribuisce notevolmente a spiegare



nei termini più adeguati la sostanza della libertà di coscienza e della libertà di religione e del fenomeno di individualizzazione della società moderna occidentale.

Il libro di Raffella Di Marzio offre al riguardo un punto di osservazione non esclusivo, ma indispensabile se si vogliono cogliere taluni aspetti specifici ed estremamente multifattoriali di tale processo.