

INTELLIGENZA ARTIFICIALE E UMANESIMO DI FRATERNITÀ

Parole chiave: Intelligenza artificiale, ChatGPT, Papa Francesco, Francescanesimo, Etica

Keywords: Artificial intelligence, ChatGPT, Pope Francis, Franciscanism, Ethics

Schlüsselwörter: Künstliche Intelligenz, ChatGPT, Papst Franziskus, Franziskanertum, Ethik

Słowa kluczowe: sztuczna inteligencja, ChatGPT, papież Franciszek, franciszkanizm, etyka

L'intelligenza artificiale e, in generale, le tecnologie dell'informazione e della comunicazione (TIC) fanno ormai parte della nostra vita¹. Non sono semplici strumenti, ma un ambiente di vita. Come possiamo abitare responsabilmente questo nuovo ambiente culturale in cui siamo tutti immersi? Come orientare correttamente il suo grande potenziale verso il bene ed evitare, al contempo, i gravi rischi che comporta?

Rispondendo al crescente interesse che questo argomento suscita nella società e nella Chiesa, nel gennaio 2024, il Papa ha dedicato due importanti messaggi all'IA, quello per la Giornata Mondiale della Pace² e quello per la Giornata delle Comunica-

* Martín Carbajo-Núñez, OFM, è nato a Figueruela de Arriba (Zamora, Spagna). Ha conseguito il dottorato in teologia morale (Alfonsianum, Roma), la laurea in filologia germanica (Università di Santiago de Compostela), il master in comunicazioni sociali (Pontificia Università Gregoriana, Roma) ed è tecnico informatico aziendale. Attualmente insegna teologia morale ed etica della comunicazione in tre università: due a Roma (Italia): Pontificia Università *Antonianum* (PUA) e Pontificia Università Lateranense *Alfonsiana* (PUL); una negli Stati Uniti: la FST, affiliata all'Università di San Diego (California). Alla PUA è stato vicerettore e rettore *Magnificus ad interim*, ORCID: <https://orcid.org/0000-0002-2814-5688>, e-mail: mcarbajo@fst.edu

¹ Alcune abbreviazioni: IA: Intelligenza Artificiale ("AI", in inglese); LLM = *Large Language Model* (Modello linguistico di grandi dimensioni). GPT = *Generative Pretrained Transformer* (Trasformatore generativo pre-addestrato). NLP: *Natural language processing* (Elaborazione del linguaggio naturale); RLHF: *Reinforcement learning from human feedback* (Apprendimento per rinforzo a partire dal feedback umano). Si tratta di una tecnica di apprendimento automatico (*Machine learning* = ML). La versione in inglese di questo articolo verrà pubblicata in „*Carthaginensia*” 79(2025).

² Francesco, *Messaggio per la 57 Giornata Mondiale della Pace* (1.01.2024), OsRom 286(2023), p. 2-3 (di seguito abbreviato: GMP).

zioni Sociali³. In precedenza, la Pontificia Accademia per la Vita aveva organizzato conferenze e simposi sulle implicazioni dei nuovi progressi tecnici nel campo della salute⁴. Il 10 gennaio 2023, la stessa Accademia Pontificia ha facilitato la partecipazione di rappresentanti dell'ebraismo, dell'islam e di altre religioni alla revisione del documento “*Rome Call for AI Ethics*”, che era stato firmato nel 2020 da aziende come Microsoft e IBM, tra le altre⁵.

In questo articolo, vengono esaminate alcune sfide etiche poste dalle tecnologie di intelligenza artificiale alla luce del magistero recente di Papa Francesco. Nella prima parte, si accenna alla difficoltà di definire questo concetto e al salto qualitativo dell'IA generativa. Questi rapidi progressi offrono “entusiasmani opportunità e gravi rischi”, che devono essere affrontati a livello legislativo ed etico (seconda parte). Il Papa esorta ad adottare, come principio etico regolatore, un umanesimo di fraternità ispirato a Francesco d'Assisi (terza parte).

1. L'IA: CONCETTO E RAPIDO SVILUPPO

Si parla spesso di quattro rivoluzioni industriali⁶: la prima favorita dall'invenzione della macchina a vapore (XVIII secolo), la seconda dall'elettricità (fine del XIX secolo), la terza dalla tecnologia digitale (metà del XX secolo). L'intelligenza artificiale avrebbe avviato la quarta, caratterizzata dalla capacità di predire le attività umane e dalla convergenza di tecnologie come le nanotecnologie, le biotecnologie e la robotica. Più che una nuova rivoluzione, sarebbe un'evoluzione dell'automazione iniziata nel XVIII secolo, anche se i suoi effetti sociali potrebbero essere rivoluzionari.

Nella prima rivoluzione industriale, le macchine hanno iniziato a svolgere molte attività manuali, che richiedevano forza muscolare piuttosto che mentale, soppiantando così i lavoratori meno qualificati (“colletti blu”). Al contrario, ora l'IA sta assumendo compiti che richiedono specializzazione e capacità cognitive, sostituendo molti impiegati qualificati della classe media (“colletti bianchi”)⁷. Tuttavia, con l'IA risulta

³ Francesco, *Messaggio per la 58 Giornata Mondiale delle Comunicazioni Sociali* (24.01.2024), OsRom 19(2024), p. 8 (di seguito abbreviato: GCS).

⁴ V. Paglia, R. Pecoraro (edd.), *Robo-Ethics. Humans, machines and health*, Pontifical Academy for Life: Vatican City 2020; V. Paglia, R. Pecoraro (edd.), *The 'good' algorithm? Artificial intelligence: ethics, law, health*, Pontifical Academy for Life: Vatican City 2021.

⁵ *AI ethics: an Abrahamic commitment to the Rome call* (10.01.2023), in: *RenAssance Foundation* [on-line], <https://www.romecall.org/ai-ethics-an-abrahamic-commitment-to-the-rome-call-2/> (Accesso: 10.03.2024).

⁶ Cf. K. Schwab, *The Fourth Industrial Revolution*, Crown Business: New York 2017; J. Rifkin, *The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World*, Palgrave Macmillan: New York 2011. Sulla prima e seconda rivoluzione industriale guidata dall'energia a vapore e dall'elettricità: J. Mokyr, *The Lever of Riches: Technological Creativity and Economic Progress*, Oxford University Press: New York 1990.

⁷ A. Gisotti, *Entrevista a Benanti: La inteligencia artificial al servicio del bien común* (12.12.2023), in: *Vatican News* [on-line], (<https://www.vaticannews.va/es/mundo/news/2023-12/benanti-la-inteligencia-artificial-al-servicio-del-bien-comun.html>) (10.07.2024).

difficile imitare le nostre capacità sensoriomotorie e percettive. Non sarà facile, ad esempio, per l'IA fare il lavoro di un idraulico⁸.

1.1. UN CONCETTO COMPLESSO E DIFFICILE DA DEFINIRE

L'IA è tra noi da molto tempo in varie forme: reti sociali, assistenti virtuali, pagamenti elettronici, motori di ricerca, traduzione automatica, riconoscimento facciale e vocale, robot, droni, auto a guida autonoma, ecc. F. Patsch indica quattro tappe fondamentali nella sua implementazione: nel 2010, si è iniziato a usarla per migliorare la qualità delle ricerche su Internet e fornire risultati più pertinenti; nel 2014, per “leggere la mente” dell'utente e anticipare ciò che sta cercando, aiutandolo con assistenti virtuali come Cortana e Alexa⁹; nel 2018, per vedere e riconoscere gli oggetti (IA di percezione); nel 2022, infine, è stata introdotta l'IA generativa¹⁰.

Il termine “IA” si applica a “una galassia di realtà diverse”, per cui è molto difficile darne una definizione univoca¹¹. “Abbraccia una varietà di scienze, teorie e tecniche volte a far sì che le macchine riproducano o imitino” le capacità associate all'intelligenza umana, come la creatività, l'apprendimento e la pianificazione¹². Forse sarebbe più conveniente usare il termine al plurale (“intelligenze artificiali”), perché i loro compiti sono frammentari e imitano o riproducono solo parzialmente alcune capacità umane. Inoltre, una volta progettate, il loro impatto dipende anche da come vengono utilizzate.

L'assenza di una definizione univoca complica il lavoro di coloro che cercano di regolamentarne l'uso e, allo stesso tempo, fa sì che alcune aziende possano aggirare i limiti imposti da eventuali normative.

⁸ “It is comparatively easy to make computers exhibit adult-level performance in solving problems on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility”: H. Moravec, *Mind children. The future of robot and human intelligence*» Harvard UP: Cambridge 1988, p. 15.

⁹ RankBrain, introduced in 2015, “uses artificial intelligence techniques to process and interpret search queries, with the goal of providing more relevant search results to users”: M. Johnsen, *The future of Artificial intelligence in digital marketing. The next big technological break*, CreateSpace: US 2021, p. 95. L'assistente virtuale Siri è stato introdotto da Apple nell'ottobre 2011. È stato seguito da molti altri.

¹⁰ F. Ferenc, *L'Intelligenza artificiale generativa e il nostro futuro. Una urgente necessità di regolamentazione*, “La Civiltà Cattolica” 4162/4(2023), p. 313-325, nota 1; J. Holmström, *From AI to digital transformation. The AI readiness framework*, „Business Horizons” 65/3(2022), p. 329-339.

¹¹ Il termine IA è stato coniato alla Conferenza di Dartmouth presso l'omonimo College di Hannover, nel New Hampshire, nell'estate del 1956. Russell e Norving indicano quattro tipi di definizioni di IA. Cf. S.J. Russell, P. Norving, *Artificial Intelligence: A Modern Approach*, 3rd ed., Prentice Hall: Upper Saddle River (NJ) 2010; In generale, possiamo definire l'IA come “qualsiasi dispositivo o programma informatico in grado di svolgere compiti normalmente associati a esseri intelligenti, come l'apprendimento, il ragionamento, la percezione, la soluzione dei problemi, il linguaggio”. Sulla sua evoluzione storica: N.J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge University Press: Cambridge 2010.

¹² Cf. M.L. Santaella Braga (ed.), *Inteligência artificial & redes sociais*, EDUC: São Paulo, 2019; K.-F. Lee, É. Roy, *I.A. la plus grande mutation de l'histoire : Qui dominera l'IA. dominera le monde*, J'ai lu: Paris, 2021.