

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 “entusiasmanti 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, E. Roy, *I.A. la plus grande mutation de l'histoire : Qui dominera l'I.A. dominera le monde*, J'ai lu: Paris, 2021.

1.2. UN SALTO QUALITATIVO: L'IA GENERATIVA

Il 30 novembre 2022, si è verificato un salto qualitativo con il lancio di ChatGPT, un'applicazione di IA generativa, capace di “produrre testi sintatticamente e semanticamente coerenti”¹³.

ChatGPT è un'applicazione di conversazione (*chat*) basata su un modello linguistico di grandi dimensioni (LLM) del tipo “GPT”. “Non è progettata per dare risposte precise e accurate”, ma per conversare e intrattenere l'interlocutore. Gestisce bene il linguaggio e produce testi belli, ma non necessariamente oggettivi e affidabili. Utilizza algoritmi avanzati di apprendimento automatico (*Deep learning*) “basati sul feedback umano” (RLHF) ed è in grado di sviluppare “comportamenti e capacità non esplicitamente programmati dagli sviluppatori”¹⁴.

Ben presto sono arrivate sul mercato numerose applicazioni simili. Ad esempio, quelle in grado di generare testo (Copilot, Gemini, Perplexity), creare immagini (Dall-E, Midjourney, Leonardo, Jasper) e produrre audio e video (Synthesia, Fliki, Pictory), per citare solo alcuni esempi.

Questa rapida evoluzione “sta modificando in modo radicale anche l'informazione e la comunicazione e, attraverso di esse, alcune basi della convivenza civile” (GCS 2024). In effetti, diventa sempre più difficile distinguere tra informatica e pensiero. In questo nuovo contesto culturale, come possiamo raggiungere la saggezza e la “comunicazione pienamente umana”? Possiamo sentirci sicuri con sistemi di intelligenza artificiale capaci di riprogrammarsi autonomamente?

2. ENTUSIASMANTI OPPORTUNITÀ E GRAVI RISCHI

Le “entusiasmanti opportunità e gravi rischi” (GMP 2024,1) che comporta l'intelligenza artificiale generativa, insieme alla sua rapida adozione, hanno messo in allarme la comunità internazionale. ChatGPT è stato lanciato il 30 novembre 2022 e in appena due mesi ha raggiunto 100 milioni di utenti attivi mensili. Nel settembre 2023, ha superato 1,5 miliardi di visite mensili, diventando “l'app di consumo con la crescita più rapida della storia”¹⁵.

Il 22 marzo 2023, più di 1.000 esperti di tecnologia, ricercatori e investitori hanno firmato una lettera aperta con la quale mettevano in guardia dai “profondi rischi per la società” posti dai giganteschi sistemi di intelligenza artificiale. La lettera chiedeva di sospenderne lo sviluppo per sei mesi per valutare e affrontare meglio i rischi.

¹³ GMP 2024, 3. Cf. Z. Gacovski, *Natural Language processing*, Arcler Press: Burlington (ON) 2021.

¹⁴ N. Da Silva Gonçalves, *Intelligenze artificiali e intelligenze incarnate: quale frontiera? Intervista a Paolo Benanti*, “La Civiltà Cattolica” 4164(2023), p. 572-586.

¹⁵ „The fastest-growing consumer application in history”: V. Vara, *One Year In and ChatGPT Already Has Us Doing Its Bidding*, NYT (31.12.2023) 5, sec. SR; L'app consumer che ha registrato la crescita più rapida è stata TikTok, che ha impiegato 9 mesi per raggiungere questa cifra.

Tra i principali firmatari vi erano specialisti di alto livello come Yoshua Bengio¹⁶ e imprenditori come Elon Musk, cofondatore di OpenAI Elon Musk e il cofondatore di Apple Steve Wozniak¹⁷. In effetti, l'allarme sollevato ha portato a limitare alcune delle potenzialità dell'IA e ad "addomesticare" in parte l'espressività delle sue conversazioni¹⁸.

Stephen Hawking aveva dichiarato che "il successo nella creazione di intelligenza artificiale sarebbe l'evento più grande della storia umana. Purtroppo, potrebbe anche essere l'ultimo, a meno che non impariamo a evitare i rischi"¹⁹.

2.1. SISTEMI AUTONOMI CHE AGISCONO IN MODO PERICOLOSO

L'intelligenza artificiale è già presente in sistemi capaci di agire autonomamente in situazioni complesse e in altri che possono riprogrammarsi, adattando i mezzi che utilizzano per raggiungere meglio lo scopo per cui sono stati progettati. È il programmatore umano che ne decide la finalità, ma queste macchine possono avere la capacità di scegliere i mezzi. Sembrerebbe che, per esse, il fine giustifichi i mezzi, contraddicendo così uno dei grandi principi etici.

Ad esempio, a volte i veicoli con guida autonoma vengono pubblicizzati come se fossero in grado di prendere decisioni in circostanze che richiedono giudizi di valore, il che richiederebbe discernimento e che quindi non è alla portata di una macchina²⁰. Per venderli più facilmente, i produttori potrebbero programmarli in modo da privilegiare sempre la sicurezza del conducente, anche se questo significa mettere seriamente a rischio la vita dei pedoni. I dilemmi del tram²¹ e del casco²² sono buoni esempi delle sfide etiche che ne derivano.

¹⁶ Riconosciuto come uno dei "padri" dell'IA per il suo contributo allo sviluppo dell'apprendimento profondo e delle reti neurali. Cf. Y. Bengio, *Learning deep architectures for AI*, Now Pub.: Hanover (Mass) 2009.

¹⁷ C. Metz, G. Schmidt, *Tech leaders urge a pause in A.I., citing 'profound risks to society'*, NYT (30.03.2023) 5, sec. 5.

¹⁸ K. Roose, *My A.I. Homewrecker Has Turned Into an A.I. Office Drudge*, NYT (15.02.2024) 1, sec. A.

¹⁹ „Successful A.I. ‘would be the biggest event in human history. Unfortunately, it might also be the last’“: N. Bilton, *Artificial Intelligence as a Threat*, NYT (6.11.2014) 2, sec. E; „Humans, who are limited by slow biological evolution, couldn't compete, and would be superseded“: R. Cellan-Jones, *Stephen Hawking warns artificial intelligence could end mankind*, (2.12.2015), in: *BBC News* [on-line], <https://www.bbc.com/news/technology-30290540> (Accesso: 10.07.2024).

²⁰ It is not possible to embed algorithms into those vehicles to “automate complex ethical decision making”: J. Millar, *Ethics settings for autonomous vehicles*, in: ed. P. Lin, R. Jenkins, K. Abney, *Robot ethics 2.0. From autonomous cars to artificial intelligence*, Oxford UP: New York 2020, p. 20-34, qui 22.

²¹ This dilemma was developed by philosopher Philippa Foot in 1967 and adapted by Judith Jarvis Thomson in 1985. Cf. L. D'Olimpio, *The trolley dilemma: would you kill one person to save five?* (3.06.2016), in: *The Conversation* [on-line], <https://theconversation.com/the-trolley-dilemma-would-you-kill-one-person-to-save-five-57111> (Accesso: 10.07.2024).

²² Cf. P. Lin, *The ethical dilemma of self-driving cars*, in: *YouTube* [on-line], <https://www.youtube.com/watch?v=ixIoDYVfKA0&t=8s> (11.07.2024); J. Markoff, *Should Your Driverless Car Hit a Pedestrian if It Would Save Your Life?*, NYT (24.06.2016) 2, sec. B.; A. Shariff, L. Rahwan, J-F. Bonnefon, *Whose life should your car save?*, NYT (6.11.2016) 9, sec. SR.

Ancora più preoccupanti sono i robot militari autonomi (LAWS: *Lethal Autonomous Weapon Systems*²³), che potrebbero superare le prestazioni per cui sono stati programmati²⁴. È stato riferito che il 27 marzo 2020 un drone turco (STM Kargu-2) ha deciso autonomamente di attaccare le truppe nemiche di Haftar nel conflitto armato in Libia²⁵, senza attendere alcun intervento umano e andando oltre quanto era prevedibile²⁶. Due anni prima, il Parlamento dell'Unione Europea aveva già adottato una risoluzione (2018/2752) che chiede la messa al bando dei LAWS a livello internazionale²⁷.

Nel maggio 2023, presso la Royal Aeronautical Society di Londra, il colonnello dell'aeronautica statunitense Tucker "Cinco" Hamilton ha rivelato che, in una simulazione, un drone gestito dall'IA aveva ucciso l'operatore perché gli aveva impedito di raggiungere i suoi obiettivi. La notizia ha suscitato grande scalpore in tutto il mondo ed è stata rapidamente smentita²⁸, ma ha fatto emergere la paura latente che queste macchine suscitano nell'immaginario collettivo.

Il cinema ha esplorato ampiamente la temuta ribellione delle macchine. Uno dei film classici è "2001: Odissea nello spazio", diretto da Stanley Kubrick e uscito nell'aprile del 1968. In esso, il computer HAL 9000 si ribella contro l'equipaggio della missione spaziale. Argomenti dello stesso genere sono stati affrontati successivamente da molti altri film, ad esempio *Blade Runner* (1982), *War Games* (1983), *RoboCop* (1987), *Terminator* (1984), *AI* (2001), *Matrix* (1999), *I Robot* (2004), *Her* (2013), ecc.

²³ The LAWS are often called "killer robots": J-B.J. Vilmer, *Terminator ethics: Should we ban 'killer robots'?*, in: Ph. Otto, E. Gräf (ed.), *3th1cs. A reinvention of ethics in the digital age?*, iRights media: Berlin 2017, p. 98-113, qui 98. In uno scenario fantascientifico, si potrebbe pensare a LAWS completamente autonome, progettate e programmate da altre macchine. Cf. J-B.J. Vilmer, *Terminator ethics*, p. 100.

²⁴ "They may at some point perform unforeseen actions, escaping the predicted area of evolution and thus contradicting the objective set by a responsible human agent": I. Jurkovic (archbishop), *Statement to the 2017 group of Governmental experts on LAWS* (Geneva 13.11.2017), in: *Permanent Observer Mission of the Holy See in Geneva* [on-line], <https://nuntiusge.org/wp-content/uploads/2020/05/20171113.pdf> (Accesso: 11.07.2024).

²⁵ M. Cramer, *A.I. Drone May Have Acted on Its Own in Attack in Libya*, NYT (4.06.2021) 8, sec. A.

²⁶ D. Lambert, *Autonomous weapons and cyberconflicts. How Christian ethics can deal with war's new face?*, "Angelicum" 97/1(2020), p. 93-105, qui 95.

²⁷ H. Blix, *A farewell to wars. The growing restraints on the interstate use of force*, Cambridge UP: Cambridge 2023, 128; cf. *European Parliament resolution of 12 September 2018 on autonomous weapon systems*, in: *European Parliament* [on-line], https://www.europarl.europa.eu/doceo/document/TA-8-2018-0341_EN.html (Accesso: 10.07.2024).

²⁸ Ch.R. Davis, P. Squire, *Air force colonel backtracks over his warning about how AI could go rogue and kill its human operators* (3.07.2023), in: *The Business Insider* [on-line], <https://www.businessinsider.com/ai-powered-drone-tried-killing-its-operator-in-military-simulation-2023-6?r=US&IR=T> (Accesso: 10.07.2024).

2.2. CONTAMINAZIONE COGNITIVA E ALTERAZIONE DELLA REALTÀ

“Chiunque sia il leader dell’IA sarà il leader del mondo”, dichiarava Vladimir Putin già nel 2017²⁹, poco dopo che il suo Paese era stato accusato di aver interferito nelle elezioni presidenziali statunitensi del 2016³⁰. L’IA sta svolgendo un ruolo fondamentale nelle guerre in Ucraina³¹ e a Gaza³². Anche la NATO sta ridisegnando la propria strategia militare per raggiungere la superiorità cognitiva. In questo modo, il campo di battaglia si sta spostando verso la mente umana. Non si tratta solo di sapere di più e meglio dell’avversario³³, ma anche di manipolare la sua mente e distruggere la coesione sociale del suo Paese.

Anche in passato si cercava di intimidire, screditare e imporre la propria cultura, lingua e religione al nemico. Ora, con l’IA, è più facile bombardarlo con informazioni false ma credibili e persuasive. In effetti, l’IA può essere uno strumento formidabile di “inquinamento cognitivo, di alterazione della realtà” (GCS 2024). L’uso che criminali e terroristi possono farne risulta molto preoccupante³⁴.

In definitiva, non sarebbe corretto dire che i media e l’IA, di per sé, siano “strumenti inerti il cui valore dipende dall’uso che ne viene fatto”³⁵. Di fatto, la loro neutralità è solo apparente (GCS 2014), poiché hanno dimensioni etiche e politiche³⁶ strettamente legate agli obiettivi e agli interessi dei loro proprietari nonché al modo in cui gli sviluppatori scelgono gli algoritmi, li alimentano con dati e progettano il modo in cui interagiscono con l’utente.

²⁹ “Putin stated: whoever becomes the leader in artificial intelligence will become the ruler of the world”: D. Schoeni, T. Vestner, *Ethical Dilemmas in the Global Defense Industry*, Oxford UP: New York 2023, p. 128.

³⁰ K.H. Jamieson, *Cyberwar. How Russian hackers and trolls helped elect a president. What we don’t, can’t, and do know*, Oxford UP: Oxford 2020; E.V.W. Davis, *Shadow warfare. Cyberwar policy in the United States, Russia, and China*, Rowman & Littlefield: London 2021.

³¹ Cf. J. Wagstaff, *Un nuevo modelo de ejército*, in: *Fondo Monetario Internacional* [on-line], <https://www.imf.org/es/Publications/fandd/issues/2023/12/Case-Studies-New-model-army-Jeremy-Wagstaff> (Acceso: 7.03.2024).

³² The Israeli military says it’s using an AI system, named ‘the Gospel,’ to select many of these targets in real-time. Cf. N. Stehr, *Understanding society and Knowledge*, Edward Elgar: Northampton (Mass) 2023, p. 41; H. Davies, B. McKernan, D. Sabbagh, ‘The Gospel’: how Israel uses AI to select bombing targets in Gaza, in: *The Guardian* [on-line], <https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets> (Acceso: 11.07.2024).

³³ „Cognitive superiority, in a NATO context, could thus be defined as an ability to excel in understanding and decision-making”: K. Paulauskas, *Why cognitive superiority is an imperative* (6.02.2024), in: *NATO Review* [on-line], <https://www.nato.int/docu/review/articles/2024/02/06/why-cognitive-superiority-is-an-imperative/index.html> (11.07.2024); Cf. D.S. Hartley, K.O. Jobson, *Cognitive superiority. Information to Power*, Springer: Cham 2021, p. 223-228.

³⁴ Cf. A.K. Cronin, *Power to the People: How Open Technological Innovation is Arming Tomorrow’s Terrorists*, Oxford UP: New York 2020.

³⁵ Cf. Pontificio Consiglio delle Comunicazioni Sociali, *Appello agli ordini contemplativi* (3.06.1973), in: *The Holy See*, https://www.vatican.va/roman_curia/pontifical_councils/pccs/documents/rc_pc_pccs_doc_03061973_contemplative-religious_it.html (Acceso: 6.02.2024).

³⁶ L. Mumford, *Authoritarian and Democratic Technics*, „Technology and Culture” 5/1(1964), p. 1-8.

“Ogni tecnologia porta con sé un mutamento del rapporto con il mondo, una facilitazione di certi aspetti di quel rapporto e una complicazione di altri. Ciò è precisamente non neutrale, dal momento che tutto dipende da quali aspetti della vita sono facilitati e quali ostacolati³⁷.”

Lo stesso accade con altre tecnologie, considerate di per sé neutrali, ma che in realtà rispondono a interessi particolari e provocano conseguenze imprevedute per il modo in cui sono state progettate e applicate. Ciò si verifica anche nell’architettura e nell’ordinamento urbano. Ad esempio, Lewis Mumford indica che molti dei ponti sulle strade di Long Island, a New York, sono stati progettati molto bassi in modo che i bus non potessero passare, tenendo così lontani da quella zona i poveri e i neri che solitamente li usano come mezzo di trasporto³⁸.

L’enciclica *Laudato si’* aveva già affermato che i media rispondono agli interessi commerciali delle multinazionali che li controllano e, quindi, promuovono lo stesso modello di sviluppo capitalistico che ha provocato l’attuale crisi socio-ambientale³⁹. Essi condizionano non solo il nostro stile di vita, ma anche il nostro orizzonte simbolico e il mondo in cui viviamo.

2.3. TUTTO È RIDOTTO A SEMPLICI DATI

Le tecnologie digitali tendono a ridurre la realtà a ciò che un computer può elaborare, tralasciando le dimensioni più fondamentali dell’esperienza umana. Secondo Jean Baudrillard, i nuovi media stanno creando soggetti autistici, perché abbandonano il simbolico e lo sostituiscono con segni e immagini autoreferenziali⁴⁰. Peter Brooks aggiunge che “la metafora è un linguaggio di fronte al quale il computer è sordo. Senza metafora invece muti sono i profeti, i sacerdoti, i teologi⁴¹”.

“Grandi possibilità di bene accompagnano il rischio che tutto si trasformi in un calcolo astratto, che riduce le persone a dati, il pensiero a uno schema, l’esperienza a un caso, il bene al profitto, e soprattutto che si finisca col negare l’unicità di ogni persona e della sua storia, col dissolvere la concretezza della realtà in una serie di dati statistici” (GCS 2024).

³⁷ J. Lynch, *Il profumo dei limoni. Tecnologia e rapporti umani nell’era di Facebook*, Lindau: Torino 2012, p. 57.

³⁸ It is crucial to pay attention “to the meaning of the designs and arrangements of our artifacts”. Cf. L. Winner, *Do Artifacts Have Politics?*, „Daedalus” 109/01(1980), p. 121-136, qui 125, <https://www.jstor.org/stable/20024652>. Indeed, many of the overpasses on Long Island, New York, were deliberately designed as extraordinarily low to keep out poor people and blacks, who normally use buses. Cf. L. Winner, *Do Artifacts*, p. 123-124.

³⁹ Cf. Francesco, *Laudato si’*. *Lettera enciclica* (24.05.2015), AAS 107(2015), n. 47.

⁴⁰ J. Baudrillard, *Il delitto perfetto. La televisione ha ucciso la realtà?*, Raffaello Cortina, Milano 1996, p. 153; Sull’enorme impatto dell’IA sulla società: L. Luciano, *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, Oxford University Press: Oxford 2014; K. Easa, *Mans Relationships with Technology An Analysis of Marshall McLuhan’s “Understanding Media. The Extensions of Man”*, Ebook, GRIN: München 2019.

⁴¹ P. Brooks, *La comunicazione della fede nell’età dei media elettronici*, ElleDiCi: Leumann 1987, p. 44.

L'uso massiccio delle tecnologie dell'IA può trasformarle in arbitri della verità, che decidono ciò che è importante e condizionano il nostro modo di pensarci⁴².

Se usata male, l'IA può costituire una minaccia molto seria per la dignità della persona umana. Infatti, elaborando tutto freddamente, come un codice binario, l'IA può mettere a rischio la nostra ammirazione per il mistero della persona, esacerbare l'individualismo egoistico e “mettere da parte i valori essenziali della compassione, della misericordia e del perdono” (GMP 2024, 5). Avendo perso il rispetto per la privacy personale, “gli errori sistemici possono facilmente moltiplicarsi”, esponendo le persone “a forme di pregiudizio e discriminazione” (*Ibid.*).

Il Papa avverte che, a nostra insaputa, le tecnologie dell'IA estraiono e strutturano i dati a fini commerciali o politici, limitando così il nostro “consapevole esercizio della libertà di scelta” (GMP 2024, 2). Una volta inserite le informazioni personali nei loro giganteschi database, possono “eliminare la possibilità che un individuo cambi e si lasci alle spalle il passato” (GMP 2024, 5).

Il pericolo per la privacy e la pace sociale è accentuato dal fatto che gli algoritmi autodidatti dell'IA, che guidano i social media, danno la priorità ai messaggi provocatori, polemici e polarizzanti, perché ciò aumenta l'audience e genera maggiori profitti. Favoriscono anche le false notizie (*fake news*), semplicistiche e dirette, a scapito della verità, che in genere risulta più complessa e meno accattivante⁴³.

2.4. NECESSITÀ DELL'ETICA

Papa Francesco fa notare che “non possiamo presumere a priori” che lo sviluppo dell'IA “apporti un contributo benefico al futuro dell'umanità”. Sarà positivo solo se ne faremo un uso responsabile e rispetteremo “i valori umani fondamentali”, a cominciare dalla “dignità intrinseca di ogni persona e dalla fraternità” (GMP 2024, 2).

L'approccio del Papa all'IA è critico, ma non pessimistico. Invita a evitare “letture catastrofiche” e, allo stesso tempo, esorta “la Comunità delle nazioni a lavorare unita al fine di adottare un trattato internazionale vincolante, che regoli lo sviluppo e l'uso dell'intelligenza artificiale nelle sue molteplici forme”⁴⁴. Ad esempio, è chiaro

⁴² Cf. V. Eubanks, *Automating inequality: How high-tech tools profile, police and punish the Poor*, St Martin's Press: New York 2018, p. 11

⁴³ Twitter, Facebook e i social media non sono progettati per una conversazione ponderata. Più il post è incendiario, più le piattaforme lo diffondono. Cf. M. Fisher, *The Chaos Machine: The inside story of how Social Media rewired our minds and our world*, Quercus: London 2023; Z. Tufekci, *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, Yale UP: New Haven, 2017; E. Pariser, *The Filter Bubble: What the Internet Is Hiding from You*, Penguin Press: New York 2011. Il documentario di Netflix «The Social Dilemma» (diretto da Jeff Orlowski, Exposure Labs, 2020) sottolinea come i social network siano progettati per coinvolgere lo spettatore attraverso contenuti sensazionalistici e incendiari, piuttosto che promuovere il dibattito sereno e ponderato.

⁴⁴ GCS 2024. Cf. OECD, *Recommendation of the Council on AI* (22.05.2019), in: *OECD Legal Instruments* [on-line]: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (Accesso: 11.07.2024); European Commission, *Ethics guidelines for trustworthy AI* (8.04.2019), in: *European Commission* [on-line], <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (10.07.2024), 9.

che l'intelligenza artificiale mette a rischio “la riservatezza, il possesso dei dati e la proprietà intellettuale” (GMP 2024, 3).

Inoltre, insiste sul fatto che, da sola, “la regolamentazione non basta” e che deve essere accompagnata da una buona formazione etica e antropologica. Senza etica, la nostra epoca rischia di essere “ricca di tecnica e povera di umanità” (GCS 2024). Pertanto, è necessario “istituire organismi incaricati di esaminare le questioni etiche emergenti” (GMP 2024, 3) e il rispetto dei diritti umani.

3. UMANESIMO DI FRATERNITÀ

Di fronte al paradigma tecnocratico e agli algoritmi che cercano la provocazione e lo scontro, il Papa propone un umanesimo di fraternità ispirato a Francesco d'Assisi. Infatti, “La dignità intrinseca di ogni persona e la fraternità che ci lega come membri dell'unica famiglia umana devono stare alla base dello sviluppo di nuove tecnologie” (GMP 2024, 2).

In un mondo in cui prevale “l'indifferenza verso il prossimo”⁴⁵, solo la promozione della fraternità e della solidarietà universale può condurre “a creare una cultura diversa, che ci orienti a superare le inimicizie e a prenderci cura gli uni degli altri”⁴⁶.

Il Concilio Vaticano II aveva già insegnato che la persona umana “è e deve essere principio, soggetto e fine di tutte le istituzioni sociali”⁴⁷. Si tratta di un essere intrinsecamente sociale, che non può “ritrovarsi pienamente se non attraverso un dono sincero di sé”⁴⁸. Pertanto, la dignità umana, la fraternità e la cura dell'ambiente sono i pilastri dell'umanesimo che deve guidare lo sviluppo dell'IA.

Questo umanesimo di fraternità, che abbraccia le creature come sorelle, trova la sua migliore espressione nel Cantico delle Creature di San Francesco. Il Santo di Assisi riconosce che “nullu homo ène dignu Te mentovare”⁴⁹. Così, “con” tutte le creature (*Cant 3: FF 263*) e “per” mezzo di esse (*Cant 5-9: FF 263*), Francesco loda il Creatore.

3.1. IA E SVILUPPO INTEGRALE

Nella nostra società materialista, lo sviluppo tecno-scientifico viene spesso equiparato al progresso integrale, dimenticando che quest'ultimo si realizza solo quando i progressi tecnici favoriscono la libertà, l'equità, la felicità e la comunione fraterna. Lo sviluppo autentico è sempre legato al bene comune, cosa che non sempre accade con le innovazioni del progresso tecnologico. Ad esempio, non si può dire che l'innovazione della bomba atomica abbia favorito il bene comune.

⁴⁵ Francesco, *Discorso al corpo diplomatico accreditato presso la Santa Sede* (11.01.2016), OsRom 7(2016), p. 4-5.

⁴⁶ Francesco, *Fratelli tutti. Lettera enciclica* (3.10.2020), AAS 112(2020), n. 57 (di seguito abbreviato: FT).

⁴⁷ Concilio Vaticano II, *Gaudium et spes. Costituzione pastorale* (7.12.1965), AAS 58(1966), n. 25.

⁴⁸ Concilio Vaticano II, *Gaudium et spes. Costituzione pastorale*, n. 24.

⁴⁹ FRANCESCO D'ASSISI, «Cantico di frate Sole» [*Cant*], n. 2, in *Fonti Francescane*, [FF], Ed. Francescane, Padova et al. ³2011, qui 263.

Lo sviluppo, per essere integrale⁵⁰, deve comprendere tre dimensioni fondamentali: materiale, sociale e spirituale (o espressiva). La prima si riferisce alla ricchezza materiale o acquisitiva (“ben-avere”), solitamente misurata dal Prodotto interno lordo (PIL)⁵¹.

La dimensione sociale si riferisce al livello di integrazione e solidarietà tra i diversi gruppi. Se le tecnologie di intelligenza artificiale non portano “a un miglioramento della qualità della vita di tutta l’umanità, ma al contrario aggravano le disuguaglianze e i conflitti, non potranno mai essere considerate vero progresso” (GMP 2024, 2). Per questo il Papa esorta ad “arginare i risvolti dannosi e discriminatori, socialmente ingiusti, dei sistemi di intelligenza artificiale” (GCS 2024). Ad esempio, V. Eubanks sostiene che negli Stati Uniti l’automazione dei servizi assistenziali sta aumentando le disuguaglianze⁵². L’IA potrebbe far aumentare ulteriormente questi divari sociali ed economici a livello globale, poiché è controllata da poche multinazionali a scopo di lucro.

La dimensione spirituale (o “espressiva”⁵³), infine, è misurata dal cosiddetto “indice di felicità”. Senza beni relazionali non c’è felicità. I beni economici sono pienamente tali quando garantiscono una vita buona, cioè quando rendono possibile la felicità pubblica. In questo senso, si dovrà affrontare l’impatto sociale delle tecnologie dell’IA, che stanno assorbendo posti di lavoro⁵⁴. La piena occupazione deve essere garantita in altri modi, data la sua importanza “per il benessere economico delle persone” (GMP 2024, 5).

3.2. ALTRE IMPLICAZIONI SOCIALI DELL’USO DELL’IA

La Chiesa ricorda che la scienza e la tecnologia ci aiutano a soddisfare i bisogni fondamentali e a superare altri mali che creano discordia e tensioni. Paolo VI aveva già affermato che “lo sviluppo è il nuovo nome della pace”⁵⁵.

Le tecnologie dell’IA “stanno cambiando il volto delle comunicazioni, della pubblica amministrazione, dell’istruzione, dei consumi, delle interazioni personali e di innumerevoli altri aspetti della vita quotidiana” (GMP 2024, 2). Promettono, ad esempio, “una rivoluzione nei processi di raccolta, organizzazione e verifica dei dati”, oltre a “introdurre importanti innovazioni nell’agricoltura, nell’istruzione e nella cultura” (GMP 2024, 6). D’altra parte, l’IA mette nelle mani dell’uomo alcune opportunità

⁵⁰ Sullo sviluppo integrale: E.F. Schumacher, *Small is beautiful die Rückkehr zum menschlichen Mass*, Oekom, München 2013; A. Schieffer, R. Ronnie, *Integral Development: Realising the Transformative Potential of Individuals, Organisations and Societies*, Gower: Farnham - Surrey (UK) 2014.

⁵¹ Cf. B. Aguilar González, *Paradigmas económicos y desarrollo sostenible. La economía al servicio de la conservación*, Euned: San José 2002, p. 37.

⁵² „Automated eligibility systems discourage them from claiming public resources that they need to survive and thrive [...]. Predictive models and algorithms tag them as risky investments”: V. Eubanks, *Automating inequality*, p. 11.

⁵³ Cf. Aristotle, *Nicomachean ethics*, Chicago UP: Chicago 2011.

⁵⁴ C.B. Frey, M.A. Osborne, *The Future of Employment: How Susceptible Are Jobs to Computerisation?*, „Technological Forecasting and Social Change” 114(2017), p. 254-280.

⁵⁵ Paulo VI, *Populorum progressio. Lettera enciclica* (26.03.1967), AAS 59(1967), n. 87.

che “possono rappresentare un rischio per la sopravvivenza e un pericolo per la casa comune” (GMP 2024, 1).

Il Papa insiste sulla “necessità di un dialogo interdisciplinare”, che includa le scienze umane e la teologia, per promuovere e guidare adeguatamente lo “sviluppo etico degli algoritmi – l’algor-etica” (GMP 2024, 6). Tali algoritmi devono essere trasparenti, in modo che se ne possano conoscere le fonti, gli obiettivi e i criteri con cui selezionano e agiscono. Non si può lasciare alle macchine la responsabilità di decisioni importanti e vitali, ad esempio in campo economico, politico e militare⁵⁶.

3.3. UNITÀ NELLA DIVERSITÀ

L’approccio critico e responsabile all’IA comporta, necessariamente, una riflessione rigorosa sul suo rapporto con la complessità della realtà, puntando a preservare il pluralismo e ad evitare disuguaglianze e ingiustizie. Oggi i social network favoriscono un “febrile scambio di opinioni” che nella maggior parte dei casi non è un vero dialogo, ma “solo monologhi che procedono paralleli” (FT 200), arrivando persino alla violenza verbale⁵⁷. Occorre “mettersi seduti ad ascoltare l’altro”, prestargli attenzione, come faceva San Francesco (FT 48).

Dobbiamo evitare che l’IA venga utilizzata per “omogeneizzare il mondo” (FT 52), “ostacolando il confronto tra le differenze” (FT 45). Alcuni dati sembrano confermare che stiamo andando nella direzione sbagliata. Le fonti si stanno riducendo a una sola e si sta costruendo “un pensiero unico elaborato algoritmicamente” (GCS 2024). L’IA viene già utilizzata “in campagne di disinformazione” (GMP 2024, 2-3) volte a condizionare l’opinione pubblica e ad interferire nelle elezioni⁵⁸.

Alcuni autori sostengono che l’IA può essere considerata una delle tecniche autoritarie e centralizzatrici⁵⁹, ma con il vantaggio che, “attraverso la meccanizzazione, l’automazione e la cibernetica”⁶⁰, è riuscita a neutralizzare la resistenza sociale che esse incontravano in passato. Trova così più facile imporre il pensiero unico e il controllo di massa. Le tecnologie dell’IA sono anche in grado di prevedere i nostri com-

⁵⁶ Per esempio nel campo della strategia di investimento, di quella politica e di quella militare. Cf. M. Tegmark, *Vita 3.0. Essere umani nell’era dell’intelligenza artificiale*, Raffaello Cortina: Milano 2017; M. Taddeo, L. Floridi, *How AI Can Be a Force for Good*, „Science” 361/6404 (2018), p. 751-752; B. Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World*, W.W. Norton & Company: New York 2019.

⁵⁷ FT 46; „L’abitudine di screditare rapidamente l’avversario”: FT 201; Cf. E.H. Buell, L. Sigelman, *Attack Politics: Negativity in Presidential Campaigns Since 1960*, 2nd ed., University Press of Kansas: Lawrence 2009; D.C. Mutz, *In-Your-Face Politics: The Consequences of Uncivil Media*, eBook, Princeton University Press: Princeton 2015.

⁵⁸ M. Gaggi, *Le falsità intelligenti che minacciano il mondo* (12.10.2023), in: *Corriere della Sera* [on-line], https://www.corriere.it/opinioni/23_ottobre_12/falsita-intelligenti-che-minacciano-mondo-5bd02be6-6927-11ee-9c9c-ce3429611a2d.shtml (10.07.2024).

⁵⁹ „From late neolithic times, [...] two technologies have recurrently existed side by side: one authoritarian, the other democratic, the first system-centered, immensely powerful, but inherently unstable, the other man-centered, relatively weak, but resourceful and durable”: L. Mumford, *Authoritarian and Democratic Technics*, p. 2.

⁶⁰ „Through mechanization, automation, cybernetic direction”: L. Mumford, *Authoritarian and Democratic Technics*, p. 5

portamenti, le nostre abitudini e i nostri desideri, mettendo a rischio la nostra libertà decisionale e cambiando il modo di intendere l'attività economica⁶¹. Di conseguenza, questi autori sostengono che l'IA sta minando la libertà, la giustizia e il pluralismo, portandoci rapidamente verso la dittatura del "Grande Fratello"⁶².

CONCLUSIONE

"Non abbiate paura delle nuove tecnologie"⁶³, ammoniva Giovanni Paolo II. Queste "meravigliose invenzioni"⁶⁴ sono "doni di Dio"⁶⁵. La Chiesa "si sentirebbe colpevole davanti al Signore se non li adoperasse"⁶⁶ e se non cercasse di evangelizzare la cultura che esse creano.

Lo stesso vale ora per l'IA, che presenta "entusiasmanti opportunità" (GMP 2024, 1). Citando Romano Guardini, il Papa ci esorta a "non irrigidirci di fronte al nuovo" e, allo stesso tempo, a essere sensibili al dolore e "a tutto ciò che di distruttivo e di non umano è in esso"⁶⁷.

I gravi rischi dell'IA non devono indurci a rifiutarla, ma piuttosto a promuoverne un uso responsabile. Dobbiamo "leggere e interpretare la novità del nostro tempo" (GCS 2024), collaborando con l'intera famiglia umana al processo globale di discernimento. Spetta a noi garantire che questi algoritmi promuovano una vita umana più piena. Infatti, "lo scopo e il significato" delle operazioni di IA dipendono da noi (GMP 2024, 4).

La dignità umana deve restare sempre il valore fondamentale e, pertanto, il benessere umano deve avere la priorità sul progresso tecnologico. Ciò richiede la promozione del bene comune e dei valori umani fondamentali, come "l'inclusione, la trasparenza, la sicurezza, l'equità, la riservatezza e l'affidabilità" (GMP 2024, 2).

Ognuno dovrà "decidere se diventare cibo per gli algoritmi oppure nutrire di libertà il proprio cuore, senza il quale non si cresce nella sapienza" (GCS 2024).

⁶¹ A. Agrawal, J. Gans, A. Goldfarb, *Prediction Machines. The simple economics of Artificial Intelligence*, Harvard Business Review Press: Boston 2022.

⁶² Cf. G. Orwell, *1984*, Signet Book: New York 1953.

⁶³ Giovanni Paolo II, *Il Rapido Sviluppo. Lettera apostolica* (24.01.2005), AAS 97(2005), n. 14.

⁶⁴ Concilio Vaticano II, *Inter mirifica. Decreto* (4.12.1963), AAS 56(1964), n. 1.

⁶⁵ Pontificio Consiglio delle Comunicazioni Sociali, *Communio et Progressio. Istruzione pastorale* (23.05.1971), AAS 63(1971), n. 2; cf. Pontificio Consiglio delle Comunicazioni Sociali, *Etica nelle Comunicazioni Sociali* (4.06.2000), OsRom (31.05.2000), suppl., inserto tabloid, n. 4; Giovanni Paolo II, *Redemptoris missio. Lettera enciclica* (7.12.1990), AAS 83(1991), n. 37; „una delle più provvidenziali conquiste del nostro tempo”: GCS 1973.

⁶⁶ Paulo VI, *Evangelii nuntiandi. Esortazione Apostolica* (8.12.1975), AAS 58(1976), n. 45.

⁶⁷ R. Guardini, *Letters from Lake Como*, citato in: GCS 2024.

INTELLIGENZA ARTIFICIALE E UMANESIMO DI FRATERNITÀ

SOMMARIO

Questo articolo presenta alcune delle «entusiasmanti opportunità» e alcuni «gravi rischi» dell'intelligenza artificiale alla luce del recente magistero di Papa Francesco. Non si tratta di un compito facile, perché il termine “Intelligenza Artificiale” (IA) si riferisce a una “galassia di realtà diverse” e non esiste una definizione univoca di essa. Recentemente c'è stato un salto di qualità con l'IA generativa (prima parte). Le sue sfide devono essere affrontate a livello legislativo ed etico (seconda parte). Più specificamente, il Papa sollecita l'adozione di un umanesimo di fraternità ispirato a Francesco d'Assisi (terza parte).

ARTIFICIAL INTELLIGENCE AND THE HUMANISM OF FRATERNITY

SUMMARY

This article presents some of the «exciting opportunities and grave risks» of artificial intelligence in light of the recent magisterium of Pope Francis. This is not an easy task, as the term «AI» refers to a «galaxy of different realities» and there is not a single definition of it. Recently there has been a qualitative leap with generative AI (1st part). Its challenges must be addressed at the legislative and ethical levels (2nd part). More specifically, the Pope urges the adoption of a humanism of fraternity inspired by Francis of Assisi (3rd part).

KÜNSTLICHE INTELLIGENZ UND DER HUMANISMUS DER BRÜDERLICHKEIT

ZUSAMMENFASSUNG

In diesem Artikel werden einige der “aufregenden Möglichkeiten” und einige “ernste Risiken” der künstlichen Intelligenz im Lichte des jüngsten Lehramtes von Papst Franziskus vorgestellt. Dies ist keine leichte Aufgabe, denn der Begriff “Künstliche Intelligenz” (KI) bezieht sich auf eine “Galaxie verschiedener Realitäten” und es gibt keine einheitliche Definition dafür. In jüngster Zeit hat es mit der generativen KI einen Quantensprung gegeben (erster Teil). Ihre Herausforderungen müssen auf gesetzgeberischer und ethischer Ebene angegangen werden (zweiter Teil). Der Papst plädiert insbesondere für einen Humanismus der Brüderlichkeit, der sich an Franz von Assisi orientiert (dritter Teil).

SZTUCZNA INTELIGENCJA I HUMANIZM BRATERSTWA

STRESZCZENIE

Niniejszy artykuł przedstawia niektóre z „ekscytujących możliwości” i niektóre z „poważnych zagrożeń” związanych ze sztuczną inteligencją w świetle niedawnego magisterium papieża Franciszka. Nie jest to łatwe zadanie, ponieważ termin „sztuczna inteligencja” (AI) odnosi się do „galaktyki różnych rzeczywistości” i nie ma dla niego jednej definicji. W ostatnim czasie nastąpił milowy krok w dziedzinie generatywnej sztucznej inteligencji (część pierwsza). Wyzwaniom z nią związanym należy stawić czoła na poziomie legislacyjnym i etycznym (część druga). W szczególności papież opowiada się za humanizmem braterstwa inspirowanym przez Franciszka z Asyżu (część trzecia).

BIBLIOGRAFIA

- Concilio Vaticano II, *Gaudium et spes. Costituzione pastorale* (7.12.1965), AAS 58(1966), p. 1025-1120.
- Concilio Vaticano II, *Inter mirifica. Decreto* (4.12.1963), AAS 56(1964), p. 145-157.
- Paolo VI, *Populorum progressio. Lettera enciclica* (26.03.1967), AAS 59(1967), p. 257-299.
- Paolo VI, *Evangelii nuntiandi. Esortazione Apostolica* (8.12.1975), AAS 58(1976), p. 5-76.
- Giovanni Paolo II, *Redemptoris missio. Lettera enciclica* (7.12.1990), AAS 83(1991), p. 249-340.
- Giovanni Paolo II, *Il Rapido Sviluppo. Lettera Apostolica* (24.01.2005), AAS 97(2005), p. 265-274.
- Francesco, *Laudato si'.* *Lettera enciclica* (24.05.2015), AAS 107(2015), p. 847-945.
- Francesco, *Discorso al corpo diplomatico accreditato presso la Santa Sede* (11.01.2016), OsRom 7(2016), p. 4-5.
- Francesco, *Fratelli tutti. Lettera enciclica* (3.10.2020), AAS 112(2020), p. 969-1074.
- Francesco, *Messaggio per la 58 Giornata Mondiale delle Comunicazioni Sociali* (24.01.2024), OsRom 19(2024), p. 8.
- Francesco, *Messaggio per la 57 Giornata Mondiale della Pace* (1.01.2024), OsRom 286(2023), p. 2-3.
- Pontificio Consiglio delle Comunicazioni Sociali, *Communio et Progressio. Istruzione pastorale* (23.05.1971), AAS 63(1971), p. 593-656.
- Pontificio Consiglio delle Comunicazioni Sociali, *Appello agli ordini contemplativi* (3.06.1973), in: *The Holy See*,
https://www.vatican.va/roman_curia/pontifical_councils/pccs/documents/rc_pc_pccs_doc_03061973_contemplative-religious_it.html (Accesso: 6.02.2024).
- Pontificio Consiglio delle Comunicazioni Sociali, *Etica nelle Comunicazioni Sociali* (4.06.2000), OsRom (31.05.2000), suppl., inserto tabloid, p. I-IV.

* * *

- Agrawal A., Gans J., Goldfarb A., *Prediction Machines. The simple economics of Artificial Intelligence*, Harvard Business Review Press: Boston 2022.
- Aguilar González B., *Paradigmas económicos y desarrollo sostenible. La economía al servicio de la conservación*, Euned: San José 2002.
- AI ethics: an Abrahamic commitment to the Rome call* (10.01.2023), in: *RenAIssance Foundation* [on-line], <https://www.romecall.org/ai-ethics-an-abrahamic-commitment-to-the-rome-call-2/> (Acceso: 10.03.2024).
- Aristotle, *Nicomachean ethics*, Chicago UP: Chicago 2011.
- Buell E.H., Sigelman L., *Attack Politics: Negativity in Presidential Campaigns Since 1960*, 2nd ed., Univ. Press of Kansas: Lawrence 2009.
- Baudrillard J., *Il delitto perfetto. La televisione ha ucciso la realtà?*, Milano 1996.
- Bengio Y., *Learning deep architectures for AI*, Now Pub.: Hanover (Mass) 2009.
- Bilton N., *Artificial Intelligence as a Threat*, NYT (6.11.2014) 2, sec. E.
- Blix H., *A farewell to wars. The growing restraints on the interstate use of force*, Cambridge UP: Cambridge 2023.
- Brooks P., *La comunicazione della fede nell'età dei media elettronici*, ElleDiCi: Leumann 1987.
- Cellan-Jones R., *Stephen Hawking warns artificial intelligence could end mankind*, (2.12.2015), in: *BBC News* [on-line], <https://www.bbc.com/news/technology-30290540> (Acceso: 10.07.2024).
- Cramer M., *A.I. Drone May Have Acted on Its Own in Attack in Libya*, NYT (4.06.2021) 8, sec. A.
- Cronin A.K., *Power to the People: How Open Technological Innovation is Arming Tomorrow's Terrorists*, Oxford UP: New York 2020.
- D'Olimpio L., *The trolley dilemma: would you kill one person to save five?* (3.06.2016), in: *The Conversation* [on-line], <https://theconversation.com/the-trolley-dilemma-would-you-kill-one-person-to-save-five-57111> (Acceso: 10.07.2024).
- Da Silva Gonçalves N., *Intelligenze artificiali e intelligenze incarnate: quale frontiera? Intervista a Paolo Benanti*, "La Civiltà Cattolica" 4164(2023), p. 572-586.
- Davis Ch.R., Squire P., *Air force colonel backtracks over his warning about how AI could go rogue and kill its human operators* (3.07.2023), in: *The Business Insider* [on-line], <https://www.businessinsider.com/ai-powered-drone-tried-killing-its-operator-in-military-simulation-2023-6?r=US&IR=T> (Acceso: 10.07.2024).
- Davis E.V.W., *Shadow warfare. Cyberwar policy in the United States, Russia, and China*, Rowman & Littlefield: London 2021.
- Davies H., McKernan B., Sabbagh D., *'The Gospel': how Israel uses AI to select bombing targets in Gaza*, in: *The Guardian* [on-line], <https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets> (Acceso: 11.07.2024).
- Easa K., *Mans Relationships with Technology An Analysis of Marshall McLuhan's "Understanding Media. The Extensions of Man"*, Ebook, GRIN: München 2019.
- Eubanks V., *Automating inequality: How high-tech tools profile, police and punish the Poor*, St Martin's Press: New York 2018.
- European Commission, *Ethics guidelines for trustworthy AI* (8.04.2019), in: *European Commission* [on-line], <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (Acceso: 10.07.2024).

- European Parliament resolution of 12 September 2018 on autonomous weapon systems*, in: *European Parliament* [on-line], https://www.europarl.europa.eu/doceo/document/TA-8-2018-0341_EN.html (Accesso: 10.07.2024).
- Fisher M., *The Chaos Machine: The inside story of how Social Media rewired our minds and our world*, Quercus: London 2023.
- Floridi L., *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, Oxford UP: Oxford 2014.
- Frey C.B., Osborne M.A., *The Future of Employment: How Susceptible Are Jobs to Computerisation?*, „Technological Forecasting and Social Change” 114(2017), p. 254-280.
- Gacovski Z., *Natural Language processing*, Arcler Press: Burlington (ON) 2021.
- Gaggi M., *Le falsità intelligenti che minacciano il mondo* (12.10.2023), in: *Corriere della Sera* [on-line], https://www.corriere.it/opinioni/23_ottobre_12/falsita-intelligenti-che-minacciano-mondo-5bd02be6-6927-11ee-9c9c-ce3429611a2d.shtml (Accesso: 10.07.2024).
- Gisotti A., *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>) (Accesso: 10.07.2024).
- Hartley D.S., Jobson K.O., *Cognitive superiority. Information to Power*, Springer: Cham 2021.
- Holmström J., *From AI to digital transformation: The AI readiness framework*, „Business Horizons” 65/3(2022), p. 329-339.
- Jamieson K.H., *Cyberwar: How Russian hackers and trolls helped elect a president. What we don't, can't, and do know*, Oxford UP: Oxford 2020.
- Johnsen M., *The future of Artificial intelligence in digital marketing. The next big technological break*, CreateSpace: US 2021.
- Jurkovic I. (archbishop), *Statement to the 2017 group of Governmental experts on LAWS* (Geneva 13.11.2017), in: *Permanent Observer Mission of the Holy See in Geneva* [on-line], <https://nuntiusge.org/wp-content/uploads/2020/05/20171113.pdf> (Accesso: 11.07.2024)
- Lambert D., *Autonomous weapons and cyberconflicts. How Christian ethics can deal with war's new face?*, “Angelicum” 97/1(2020), p. 93-105.
- Lee K.-F., ROY É., *I.A. la plus grande mutation de l'histoire : Qui dominera l'I.A. dominera le monde*, J'ai lu: Paris, 2021.
- Lin P., Jenkins R., Abney K. (ed.), *Robot ethics 2.0. From autonomous cars to artificial intelligence*, Oxford UP: New York 2020.
- Lin P., *The ethical dilemma of self-driving cars*, in: *YouTube* [on-line], <https://www.youtube.com/watch?v=ixIoDYVfKA0&t=8s> (Accesso: 11.07.2024);
- Lynch J., *Il profumo dei limoni. Tecnologia e rapporti umani nell'era di Facebook*, Lindau: Torino 2012².
- Markoff J., *Should Your Driverless Car Hit a Pedestrian if It Would Save Your Life?*, NYT (24.06.2016) 2, sec. B.
- Metz C., Schmidt G., *Tech leaders urge a pause in A.I., citing 'profound risks to society'*, NYT (30.03.2023) 5, sec. 5.
- Millar J., *Ethics settings for autonomous vehicles*, in: ed. P. Lin, R. Jenkins, K. Abney, *Robot ethics 2.0. From autonomous cars to artificial intelligence*, Oxford UP: New York 2020, p. 20-34.
- Mokyr J., *The Lever of Riches: Technological Creativity and Economic Progress*, Oxford UP: New York 1990.
- Moravec H., *Mind children. The future of robot and human intelligence*» Harvard UP: Cambridge 1988.

- Mutz D.C., *In-Your-Face Politics: The Consequences of Uncivil Media*, eBook, Princeton University Press: Princeton 2015.
- Mumford L., *Authoritarian and Democratic Technics*, „Technology and Culture” 5/1(1964), p. 1-8.
- Nilsson N.J., *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge UP: Cambridge 2010.
- OECD, *Recommendation of the Council on AI* (22.05.2019), in: *OECD Legal Instruments* [on-line]: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (Acceso: 11.07.2024).
- Orwell G., *1984*, Signet Book: New York 1953.
- Otto Ph., Gräf E. (ed.), *3th1cs. A reinvention of ethics in the digital age?*, iRights media: Berlin 2017.
- Paglia V., Pecoraro R. (ed.), *Robo-Ethics. Humans, machines and health*, Pontifical Academy for Life: Vatican City 2020.
- Paglia V., Pecoraro R. (ed.), *The ‘good’ algorithm? Artificial intelligence: ethics, law, health*, Pontifical Academy for Life: Vatican City 2021.
- Pariser E., *The Filter Bubble: What the Internet Is Hiding from You*, Penguin Press: New York 2011.
- Patsch F., *L’Intelligenza artificiale generativa e il nostro futuro. Una urgente necessità di regolamentazione*, “La Civiltà Cattolica” 4162/4(2023), p. 313-325.
- Paulauskas K., *Why cognitive superiority is an imperative* (6.02.2024), in: *NATO Review* [on-line], <https://www.nato.int/docu/review/articles/2024/02/06/why-cognitive-superiority-is-an-imperative/index.html> (Acceso: 11.07.2024).
- Rifkin J., *The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World*, Palgrave Macmillan: New York 2011.
- Rodríguez Herrera I., Ortega Carmona A., *Los escritos de San Francisco de Asís*, Espigas: Murcia 2003².
- Roose K., *My A.I. Homewrecker Has Turned Into an A.I. Office Drudge*, NYT (15.02.2024) 1, sec. A.
- Russell S.J., Norving P., *Artificial Intelligence: A Modern Approach*, 3rd ed., Prentice Hall: Upper Saddle River (NJ) 2010.
- Santaella Braga M.L. (ed.), *Inteligência artificial & redes sociais*, EDUC: São Paulo 2019.
- Schieffer A., Lessem R., *Integral Development: Realising the Transformative Potential of Individuals, Organisations and Societies*, Gower: Farnham - Surrey (UK) 2014.
- Schneier B., *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World*, W.W. Norton & Company: New York 2019.
- Schoeni D., Vestner T., *Ethical Dilemmas in the Global Defense Industry*, Oxford UP: New York 2023.
- Shariff A., Rahwan L., Bonnefon J-F., *Whose life should your car save?*, NYT (6.11.2016) 9, sec. SR.
- Schumacher E.F., *Small is beautiful die Rückkehr zum menschlichen Mass*, Oekom: München 2013.
- Schwab K., *The Fourth Industrial Revolution*, Crown Business: New York 2017
- Stehr N., *Understanding society and Knowledge*, Edward Elgar: Northampton (Mass) 2023.
- Taddeo M., Floridi L., *How AI Can Be a Force for Good*, „Science” 361/6404(2018), p. 751-752.

- Tegmark M., *Vita 3.0. Essere umani nell'era dell'intelligenza artificiale*, Raffaello Cortina: Milano 2017.
- Tufekci Z., *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, Yale UP: New Haven 2017.
- Vara V., *One Year In and ChatGPT Already Has Us Doing Its Bidding*, NYT (31.12.2023) 5, sec. SR.
- Vilmer J.-BJ, *Terminator ethics: Should we ban 'killer robots'?*, in: Ph. Otto, E. Gräf (ed.), *3th1cs. A reinvention of ethics in the digital age?*, iRights media: Berlin 2017, p. 98-113.
- Wagstaff J., *Un nuevo modelo de ejército*, in: *Fondo Monetario Internacional* [on-line], <https://www.imf.org/es/Publications/fandd/issues/2023/12/Case-Studies-New-model-army-Jeremy-Wagstaff> (Acceso: 11.07.2024).
- Winner L., *Do Artifacts Have Politics?*, „Daedalus” 109/01(1980), p. 121-136.