09 | Referencias

ABILab (2026). AI Banking (R)evolution: oltre la scelta. Rapporto AI Hub.

AESIA (2026). Agencia Española de Supervisión de Inteligencia Artificial. https://aesia.digital.gob.es/es

AI Act (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

AI Board (2026). Governance and coordination; AI board meetings. https://digital-strategy.ec.europa.eu/en/policies/ai-board

AICerts (2026). Generative AI Phishing Boosts Clicks, Reshapes Cyber Risk. https://www.aicerts.ai/news/generative-ai-phishing-boosts-clicks-reshapes-cyber-risk/

Altman (2025a). Three Observations. https://blog.samaltman.com/three-observations

Altman (2025b). The Gentle Singularity. https://blog.samaltman.com/the-gentle-singularity

Altman (2024a). Could AI create a one-person unicorn? Fortune. https://finance.yahoo.com/news/could-ai-create-one-person-120000722.html

Altman (2024b). The Intelligence Age. Blog personal. https://ia.samaltman.com/

Amazon (2023). Amazon announces 8 innovations to better deliver for customers, support employees, and give back to communities around the world. https://www.aboutamazon.com/news/operations/amazon-delivering-the-future-2023-announcements

Amodei (2024a). Machines of loving grace. https://www.darioamodei.com/essay/machines-of-loving-grace

Amodei (2024b). Machines of Loving Grace: How AI Could Transform the World for the Better. Blog personal. https://www.darioamodei.com/essay/machines-of-loving-grace

Amodei (2025). Technology in the World, Annual Meeting Davos 2025, World Economic Forum. https://www.weforum.org/meetings/world-economic-forum-annual-meeting-2025/sessions/technology-in-the-world/

Amodei (2026). The Adolescence of Technology. https://www.darioamodei.com/essay/the-adolescence-of-technology

Anthropic (2025). Anthropic Economic Index – September 2025 Report. https://www.anthropic.com/research/anthropic-economic-index-september-2025-report

Anthropic (2026). Claude's new constitution. Anthropic. https://www.anthropic.com/news/claude-new-constitution

Australia (2025). Australia's AI Ethics Principles. https://www.industry.gov.au/publications/australias-ai-ethics-principles

Backlinko (2025). ChatGPT / OpenAI Statistics: How Many People Use ChatGPT? https://backlinko.com/chatgpt-stats

Baker McKenzie (2025). Navigating Labor's Response to AI: Proactive Strategies for Multinational Employers Across the Atlantic. https://www.theemployerreport.com/2025/06/navigating-labors-response-to-ai-proactive-strategies-for-multinational-employers-across-the-atlantic/

Barkhuus (2003). Is Context-Aware Computing Taking Control Away from the User? Three Levels of Interactivity Examined. UbiComp 2003. Springer. https://doi.org/10.1007/978-3-540-39653-6_12

Batty (2024). Digital Twins in City Planning. *Nature Computational Science*, 4, 192–199. https://doi.org/10.1038/s43588-024-00606-5

Bettencourt (2024). Recent Achievements and Conceptual Challenges for Urban Digital Twins. Nature Computational Science, 4, 150–153. https://doi.org/10.1038/s43588-024-00604-7

Bimpas (2024). Leveraging Pervasive Computing for Ambient Intelligence: A Survey on Recent Advancements, Applications and Open Challenges. Computer Networks, 239, 110156. https://doi.org/10.1016/j.comnet.2023.110156

Bletchley Declaration (2023). The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023. https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023

Bloomberg (2026). AI Startup Nabs $100 Million to Help Firms Predict Human Behavior. https://www.bloomberg.com/news/articles/2026-02-12/ai-startup-nabs-100-million-to-help-firms-predict-human-behavior

Boston Dynamics (2025a). An Electric New Era for Atlas. https://bostondynamics.com/blog/electric-new-era-for-atlas/

Boston Dynamics (2025b). Large Behavior Models and Atlas Find New Footing. https://bostondynamics.com/blog/large-behavior-models-atlas-find-new-footing/

Brown (2025). AI's War in the Courtroom: Copyright Disputes Spike in 2025. https://www.bestlawfirms.com/articles/ai-war-in-the-courtroom-copyright-disputes-spike-in-2025/7186

Business Insider (2025a). Walmart just showed off its new AI-powered warehouses — take a look inside. https://www.businessinsider.com/see-inside-walmart-high-tech-refrigerated-grocery-warehouse-2024-7

Business Insider (2025b). The guy who coined 'vibe coding' predicts it will 'terraform software and alter job descriptions'. https://www.businessinsider.com/andrej-karpathy-coined-vibecoding-ai-prediction-2025-12

Cambridge (2025). Navigating China's regulatory approach to generative artificial intelligence and large language models. https://www.cambridge.org/core/journals/cambridge-forum-on-ai-law-and-governance/article/navigating-chinas-regulatory-approach-to-generative-artificial-intelligence-and-large-language-models/969B2055997BF42DE693B7A1A1B4E8BA

Centre for European Policy (2026). Competition in Generative AI: Updated Assessment. cepInput No. 1/2026. https://www.cep.eu/fileadmin/user_upload/cep.eu/Studien/cepInput_Competition_in_Generative_AI/cepInput_Competition_in_GenAI_Updated_Assessment.pdf

Chatgptiseatingtheworld (2026). Updated Master chart of copyright, DMCA and other claims in suits v. AI (Dec. 5, 2025). https://chatgptiseatingtheworld.com/2025/12/03/updated-master-chart-of-copyright-dmca-and-other-claims-in-suits-v-ai-dec-3-2025/

Chen (2025). Need Help? Designing Proactive AI Assistants for Programming. CHI 2025. ACM. https://doi.org/10.1145/3706598.3714002

Cheong (2025). E2E Process Automation Leveraging Generative AI and IDP-Based Automation Agent: A Case Study on Corporate Expense Processing. https://arxiv.org/abs/2505.20733

Christensen (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business School Press.

CKGSB (2025). Cheung Kong Graduate School of Business. Banking on data: How MYbank is revolutionizing supply chain finance. CKGSB Knowledge. https://english.ckgsb.edu.cn/knowledge/article/unleashing-innovation-in-china-series-banking-on-data-how-mybank-is-revolutionizing-supply-chain-finance/

Corrêa (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*, 4(10), 100857. https://doi.org/10.1016/j.patter.2023.100857

Covington (2025). New Artificial Intelligence Legislation in Mexico. Global Policy Watch. https://www.globalpolicywatch.com/2025/03/new-artificial-intelligence-legislation-in-mexico/

CrowdStrike (2025). CrowdStrike Advances Next-Gen SIEM with Threat Hunting Across Data Sources, AI-Driven UEBA. https://www.crowdstrike.com/en-us/blog/crowdstrike-advances-next-gen-siem-capabilities/

Cyberhaven (2025). AI Adoption and Risk Report Q2 2025. https://info.cyberhaven.com/hubfs/Content %20PDF/Cyberhaven %20Labs %20- %202025 %20AI %20Adoption %20& %20Risk %20Report.pdf

Darktrace (2025). New Report Finds that 78 % of Chief Information Security Officers Globally are Seeing a Significant Impact from AI-Powered Cyber Threats – up 5 % from last year. https://www.darktrace.com/news/new-report-finds-that-78-of-chief-information-security-officers-globally-are-seeing-a-significant-impact-from-ai-powered-cyber-threats

DBS Bank (2024). DBS AI-Powered Digital Transformation. https://www.dbs.com/artificial-intelligence-machine-learning/artificial-intelligence/dbs-ai-powered-digital-transformation.html

Dealroom (2025). AI startups: Revenue per employee benchmarks. https://x.com/dealroomco/status/1914264599505018989

Deutsche Bank (2025). Claudio de Sanctis, Head of Private Bank, Deutsche Bank AG Private Bank. Investor Deep Dive 2025. https://investor-relations.db.com/files/documents/other-presentations-and-events/2025/IDD-2025-Script-Private-Bank-Claudio-de-Sanctis.pdf

DHL (2024). DHL Supply Chain Passes Unprecedented 500 Million Picks Milestone Using Locus Robotics Autonomous Mobile Robots. https://www.dhl.com/es-en/home/press/press-archive/2024/dhl-supply-chain-passes-unprecedented-500-million-picks-milestone-using-locus-robotics-autonomous-mobile-robots.html

EBA (2021). EBA Discussion Paper on Machine Learning for IRB Models. https://www.eba.europa.eu/sites/default/files/document_library/Publications/Discussions/2022/Discussion %20on %20machine %20learning %20for %20IRB %20models/1023883/Discussion %20paper %20on %20machine %20learning %20for %20IRB %20models.pdf

ECB (2025). ECB Guide to Internal Models. https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf

EDPB (2025). AI Privacy Risks & Mitigations Large Language Models (LLMs). https://www.edpb.europa.eu/our-work-tools/our-documents/support-pool-experts-projects/ai-privacy-risks-mitigations-large_en

Epoch (2025a). AI Benchmarking. https://epoch.ai/benchmarks

Epoch (2025b). How much power will frontier AI training demand in 2030? https://epoch.ai/blog/power-demands-of-frontier-ai-training

ESOMAR (2024). Global Market Research 2024. https://shop.esomar.org/knowledge-center/library?publication=3019

Eurostat (2025). 32.7 % of EU people used generative AI tools in 2025. https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251216-3

Financial News London (2025). Deutsche Bank to roll out 'banking butlers' for affluent clients. https://www.fnlondon.com/articles/deutsche-bank-to-roll-out-banking-butlers-for-ultra-wealthy-clients-77e0349a

Figure (2025). F.02 Contributed to the Production of 30,000 Cars at BMW. https://www.figure.ai/news/production-at-bmw

FirstPageSage (2025). ChatGPT Usage Statistics: December 2025. https://firstpagesage.com/seo-blog/chatgpt-usage-statistics/

Fortune (2025a). Deloitte allegedly cited AI-generated research in a million-dollar report for a Canadian provincial government. https://fortune.com/2025/11/25/deloitte-caught-fabricated-ai-generated-research-million-dollar-report-canada-government/

Fortune (2025b). Elon Musk reveals massive plans for Tesla and Optimus—'Things are really going to go ballistic next year'. https://fortune.com/2025/01/30/elon-musk-reveals-massive-plans-tesla-optimus-self-driving-cars-humanoid-robots/

Fortune (2025c). AI enabled Klarna to halve its workforce—now, the CEO is warning other tech leaders to be honest about the risks. https://fortune.com/2025/10/10/klarna-ceo-sebastian-siemiatkowski-halved-workforce-says-tech-ceos-sugarcoating-ai-impact-on-jobs-mass-unemployment-warning/

Gartner (2025a). Hype Cycle for Artificial Intelligence. https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025

Gartner (2025b). Gartner Predicts Over 40 % of Agentic AI Projects Will Be Canceled by End of 2027. https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027

Google (2023). Practitioners Guide to MLOps: A framework for continuous delivery and automation of machine learning. https://cloud.google.com/resources/mlops-whitepaper

Google DeepMind (2025). AlphaFold: Science and impact. https://deepmind.google/science/alphafold/

Google Quantum AI (2024). Quantum error correction below the surface code threshold. *Nature*, 638, 920–926. https://doi.org/10.1038/s41586-024-08449-y

Grieves-Vickers (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. En Kahlen, F.J. et al. (eds.), *Transdisciplinary Perspectives on Complex Systems*. Springer. https://doi.org/10.1007/978-3-319-38756-7_4

Gu (2025). Large Language Models for Constructing and Optimizing Machine Learning Workflows: A Survey. https://arxiv.org/html/2411.10478v1

Heydari (2025). Tiny Machine Learning and On-Device Inference: A Survey of Applications, Challenges, and Future Directions. Sensors, 25(10), 3191. https://doi.org/10.3390/s25103191

HelpNetSecurity (2025). 67 % of daily security alerts overwhelm SOC analysts. https://www.helpnetsecurity.com/2023/07/20/soc-analysts-tools-effectiveness/

Hernández-Torres (2025). Challenges and Opportunities of Ambient Intelligence (AmI) in the 21st Century: A Historical Review. Evolutionary Intelligence, 18, 80. https://doi.org/10.1007/s12065-025-01010-z

Huang (2021). Power of data in quantum machine learning. *Nature Communications*, 12, 2631. https://doi.org/10.1038/s41467-021-22539-9

Hyundai (2025). Hyundai Motor Group to Unveil AI Robotics Strategy at CES 2026. https://www.hyundai.com/worldwide/en/newsroom/detail/0000001093

IBM (2025). Cost of a Data Breach Report 2025. https://www.ibm.com/reports/data-breach

IBM (2025b). Chief AI Officers cut through complexity to create new paths to value. https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/chief-ai-officer

Index Ventures (2026). Life, the Universe, and Simile: Leading Simile's $100M Series A. https://www.indexventures.com/perspectives/life-the-universe-and-simile-leading-similes-series-a/

iDanae (3Q20). Cátedra iDanae (UPM–Management Solutions). MLOps, a key element in the digital ecosystem. 2020. https://blogs.upm.es/catedra-idanae/wp-content/uploads/sites/698/2020/10/Idanae-3Q20.pdf

iDanae (2Q23). Cátedra iDanae (UPM–Management Solutions). Large Language Models: a new era for Artificial Intelligence. 2023. https://blogs.upm.es/catedra-idanae/wp-content/uploads/sites/698/2023/07/Idanae-2Q23.pdf**iDanae (1Q24).** Cátedra iDanae (UPM–Management Solutions). Towards a sustainable Artificial Intelligence. 2024. https://blogs.upm.es/catedra-idanae/wp-content/uploads/sites/698/2024/04/Idanae-1Q24-VDef.pdf

iDanae (1Q25). Cátedra iDanae (UPM–Management Solutions). The challenge of biases in the construction of Artificial Intelligence systems. 2025. https://blogs.upm.es/catedra-idanae/wp-content/uploads/sites/698/2025/04/Idanae-1Q25.pdf

iDanae (2Q25). Cátedra iDanae (UPM–Management Solutions). GenAI: an approach to multi-agents systems. 2025. https://blogs.upm.es/catedra-idanae/wp-content/uploads/sites/698/2025/07/Idanae-2Q25.pdf

IEA (2025a). Electricity 2025: Analysis and forecast to 2027. International Energy Agency. https://www.iea.org/reports/electricity-2025

IEA (2025b). World Energy Outlook Special Report on Energy and AI. International Energy Agency. https://www.iea.org/reports/energy-and-ai

IMD (2023). In the Field with Ping An. *IMD Business School*. https://www.imd.org/research-knowledge/digital/articles/in-the-field-with-ping-an/

IMF (2024). Gen-AI: Artificial Intelligence and the Future of Work. https://www.imf.org/-/media/files/publications/sdn/2024/english/sdnea2024001.pdf

ILO (2025a). International Labour Organization. Generative AI and Jobs: A Global Analysis of Potential Effects on Job Quantity and Quality. https://www.ilo.org/sites/default/files/2025-05/WP140_web.pdf

ILO (2025b). International Labour Organization. Governing AI in the World of Work: A review of global ethics guidelines. https://www.ilo.org/resource/article/governing-ai-world-work-review-global-ethics-guidelines

ILO (2025c). International Labour Organization. Global Case Studies of Social Dialogue on AI and Algorithmic Management. https://www.ilo.org/sites/default/files/2025-07/wp144_web.pdf

Inter-Parliamentary Union (2025). Parliamentary actions on AI policy. https://www.ipu.org/impact/democracy-and-strong-parliaments/artificial-intelligence/parliamentary-actions-ai-policy

Ironscales (2025). Ironscales Fall 2025 Threat Report. https://ironscales.com/hubfs/Landing %20Page %20Assets/Fall %202025 %20Threat %20Report/2025 %20Fall %20Threat %20Report_Beyond %20Detection_Reality %20of %20Deepfake %20Attacks %202.pdf

ISO/IEC (2023). ISO/IEC 42001:2023 - Artificial Intelligence Management Systems. https://www.iso.org/standard/42001

Jones (2025). Large Language Models Pass the Turing Test. https://arxiv.org/abs/2503.23674

Jumpcloud (2025). How Effective Is AI for Cybersecurity Teams? 2025 Statistics. https://jumpcloud.com/blog/how-effective-is-ai-for-cybersecurity-teams

KnowBe4 (2025). Phishing Threat Trends Report. https://www.knowbe4.com/hubfs/Phishing-Threat-Trends-2025_Report.pdf

Krijger, J. (2023). Operationalising ethics for AI in the financial industry: Insights from the Volksbank case study. *Journal of Digital Banking*, 8(3), 220–241. https://doi.org/10.69554/YQZC2796

Kumar (2020). Adversarial Machine Learning - A Taxonomy and Terminology of Attacks and Mitigations. NIST AI 100-2e2023. https://csrc.nist.gov/pubs/ai/100/2/e2023/final

Kusumegi et al. (2025). Scientific production in the era of large language models. Science. https://www.science.org/doi/10.1126/science.adw3000

Li (2025). Pitfalls and prospects of quantum machine learning. *Nature Computational Science*, 5, 1095–1097. https://doi.org/10.1038/s43588-025-00914-6

Liu (2024). Towards provably efficient quantum algorithms for large-scale machine-learning models. *Nature Communications*, 15, 434. https://doi.org/10.1038/s41467-023-43957-x

Lukac (2025). Ambient AI Scribes in Clinical Practice: A Randomized Trial. NEJM AI. https://doi.org/10.1056/AIoa2501000

Management Solutions (2023). Explainable Artificial Intelligence (XAI): Challenges of model interpretability. 2023. https://www.managementsolutions.com/en/microsites/whitepapers/explainable-artificial-intelligence

Management Solutions (3Q23). Follow-up report on Machine Learning for IRB models. Technical note on regulations, 3Q23, 2023.

https://www.managementsolutions.com/en/publications-and-events/regulatory-notes/technical-notes-on-regulations/follow-report-machine-learning-irb-models

Management Solutions (4Q24). Artificial Intelligence Act. Technical note on regulations, 4Q24, 2024. https://www.managementsolutions.com/en/publications-and-events/regulatory-notes/technical-notes-on-regulations/proposal-regulation-european-approach-artificial-intelligence

Management Solutions (4Q24a). Artificial Intelligence: regulatory landscape. Technical note on regulations, 4Q24, 2024. https://www.managementsolutions.com/en/publications-and-events/regulatory-notes/technical-notes-on-regulations/artificial-intelligence-regulatory-landscape

Marwala (2025). Why the Need for Governing Ambient Intelligence Has Never Been More Urgent. United Nations University. https://unu.edu/article/why-need-governing-ambient-intelligence-has-never-been-more-urgent

Mascelli (2025). Harvest Now Decrypt Later: Examining Post-Quantum Cryptography and the Data Privacy Risks for Distributed Ledger Networks. *Finance and Economics Discussion Series* 2025-093. Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2025.093

Microsoft (2026). AI-powered SIEM, built for modern security. https://marketingassets.microsoft.com/gdc/gdccKuCLQ/original

MIT Technology Review (2024). What's next for AlphaFold: A conversation with a Google DeepMind Nobel laureate. https://www.technologyreview.com/2025/11/24/1128322/whats-next-for-alphafold-a-conversation-with-a-google-deepmind-nobel-laureate/

MITRE (2025). ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems). https://atlas.mitre.org/

Moroni (2025). Insurmountable Limitations of City-Scale Digital Twins? On Urban Knowledge and Planning. Computational Urban Science. Springer Nature. https://doi.org/10.1007/s43762-025-00174-0

Nadella (2025). LinkedIn Post. https://www.linkedin.com/posts/satyanadella_just-wrapped-our-earnings-call-and-wanted-activity-7389433821181562880-GnMz/

Nissenbaum (1996). Accountability in a Computerized Society. Science and Engineering Ethics, 2, 25–42. https://link.springer.com/article/10.1007/BF02639315

NIST (2023). AI Risk Management Framework. https://www.nist.gov/itl/ai-risk-management-framework

NIST (2024a). Secure Software Development Practices for Generative AI and Dual-Use Foundation Models: An SSDF Community Profile. https://www.nist.gov/publications/secure-software-development-practices-generative-ai-and-dual-use-foundation-models-ssdf

NIST (2024b). Post-Quantum Cryptography Standards: FIPS 203, FIPS 204, FIPS 205. National Institute of Standards and Technology. https://csrc.nist.gov/projects/post-quantum-cryptography

Notateslaapp (2025). Tesla Eyes $20K Price Target For Optimus, Extremely Fast Production Ramp. https://www.notateslaapp.com/news/3314/tesla-eyes-20k-price-target-for-optimus-extremely-fast-production-ramp

OECD (2024). A Sectoral Taxonomy of AI Intensity. https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/12/a-sectoral-taxonomy-of-ai-intensity_c2baae71/1f6377b5-en.pdf

OECD (2025a). Emerging Divides in the Transition to Artificial Intelligence. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/emerging-divides-in-the-transition-to-artificial-intelligence_eeb5e120/7376c776-en.pdf

OECD (2025b). The effects of generative AI on productivity, innovation and entrepreneurship. OECD Artificial Intelligence Papers, no. 39. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/the-effects-of-generative-ai-on-productivity-innovation-and-entrepreneurship_da1d085d/b21df222-en.pdf

OECD (2026). AI Use by Individuals Surges Across the OECD as Adoption by Firms Continues to Expand. https://www.oecd.org/en/about/news/announcements/2026/01/ai-use-by-individuals-surges-across-the-oecd-as-adoption-by-firms-continues-to-expand.html

Ouyang (2025). FELA: A Multi-Agent Evolutionary System for Feature Engineering of Industrial Event Log Data. https://arxiv.org/html/2510.25223

OWASP (2025). OWASP Top 10 for Large Language Model Applications. https://owasp.org/www-project-top-10-for-large-language-model-applications/

Oxford (2025). AI Regulation: The Politics of Fragmentation and Regulatory Capture. https://blogs.law.ox.ac.uk/oblb/blog-post/2025/06/ai-regulation-politics-fragmentation-and-regulatory-capture

Parfit (1984). Reasons and Persons. Oxford University Press.

Park (2023). Generative Agents: Interactive Simulacra of Human Behavior. UIST '23: Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology. ACM. https://doi.org/10.1145/3586183.3606763

Park (2024). Generative Agent Simulations of 1,000 People. https://arxiv.org/abs/2411.10109

Patel (2025). U.S. AI Law & Policy Explained. Enterprise AI Governance. https://oliverpatel.substack.com/p/us-ai-law-and-policy-explained

Pew (2025). Pew Research Institute. How the U.S. Public and AI Experts View Artificial Intelligence. https://www.pewresearch.org/

Phishcare (2025). Top 10 Deepfake Phishing Scams. https://phishcare.com/top-10-deepfake-phishing-scams/

Pixiebrix (2025). Top Chief AI Officers of 2025. https://www.pixiebrix.com/reports/top-ai-officers-of-2025

PR Newswire (2023). SlashNext's 2023 State of Phishing Report Reveals a 1,265 % Increase in Phishing Emails Since the Launch of ChatGPT in November 2022, Signaling a New Era of Cybercrime Fueled by Generative AI. https://www.prnewswire.com/news-releases/slashnexts-2023-state-of-phishing-report-reveals-a-1-265-increase-in-phishing-emails-since-the-launch-of-chatgpt-in-november-2022--signaling-a-new-era-of-cybercrime-fueled-by-generative-ai-301971557.html

Proofpoint (2025). AI Threat Detection. https://www.proofpoint.com/us/threat-reference/ai-threat-detection

Pu (2025). Assistance or Disruption? Exploring and Evaluating the Design and Trade-offs of Proactive AI Programming Support. CHI 2025. ACM. https://doi.org/10.1145/3706598.3713357

Quartz (2025). AI powers smaller startups toward a new era of unicorns. https://qz.com/unicorn-entrepreneur-founder-solo-ai-startup-automation-workforce

Rawls (1971). A Theory of Justice. Harvard University Press.

Ryt Bank (2025). The World's First AI-Powered Bank. https://www.rytbank.my/

Sacra (2026). Cursor revenue, valuation & funding. https://sacra.com/c/cursor/

Shan (2024). Transitioning from MLOps to LLMOps: Navigating the Unique Challenges of Large Language Models. https://doi.org/10.3390/info16020087

Shankar (2021). Towards Observability for Machine Learning Pipelines. DOI:10.48550/arXiv.2108.13557. (PDF) Towards Observability for Machine Learning Pipelines

SoSafe (2025). Global businesses face escalating AI risk, as 87 % hit by AI cyberattacks. https://sosafe-awareness.com/company/press/global-businesses-face-escalating-ai-risk-as-87-hit-by-ai-cyberattacks/

SQ Magazine (2025). AI Cyber Attacks Statistics 2025: How Attacks, Deepfakes & Ransomware Have Escalated. https://sqmagazine.co.uk/ai-cyber-attacks-statistics/

Stanford (2023a). Stanford Encyclopedia of Philosophy. Ethics of Artificial Intelligence and Robotics. https://plato.stanford.edu/entries/ethics-ai/

Stanford (2023b). Stanford Encyclopedia of Philosophy. Philosophy of Artificial Intelligence. https://plato.stanford.edu/entries/artificial-intelligence/

Stanford (2025). The 2025 AI Index Report. https://hai.stanford.edu/ai-index/2025-ai-index-report

Stone (2025). Navigating MLOps: Insights into Maturity, Lifecycle, Tools, and Careers. https://arxiv.org/html/2503.15577v1

Tesla Car World (2025). Elon Musk Unveils Tesla Bot Gen 3 Real Homemaker Updates. https://www.youtube.com/watch?v=HR1HrrneNHs&t=1s

Teslarati (2025). Tesla Optimus' pilot line will already have an incredible annual output. https://www.teslarati.com/tesla-optimus-pilot-line-will-already-have-an-incredible-annual-output/

The Network Installers (2025). AI Cyber Threat Statistics. https://thenetworkinstallers.com/es/blog/ai-cyber-threat-statistics/

Thompson (1980). Moral Responsibility of Public Officials: The Problem of Many Hands. American Political Science Review, 74(4), 905–916. https://www.cambridge.org/core/journals/american-political-science-review/article/abs/moral-responsibility-of-public-officials-the-problem-of-many-hands/39DD3FAB7BF7DC7A242407143674F22B

Toyota Research Institute (2025). AI-Powered Robot by Boston Dynamics and Toyota Research Institute Takes a Key Step Towards General-Purpose Humanoids. https://www.tri.global/news/ai-powered-robot-boston-dynamics-and-toyota-research-institute-takes-key-step-towards-general

UK AI Safety Institute (2026). International AI Safety Report 2026. UK Government. https://www.gov.uk/government/publications/international-ai-safety-report-2026

UK DSIT (2023). UK Department for Science, Innovation and Technology. Public Attitudes to Data and AI: Tracker Survey (Report). https://www.gov.uk/government/publications/public-attitudes-to-data-and-ai-tracker-survey

UK Government (2023). A pro-innovation approach to AI regulation. Policy paper. https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

UNDP (2025). Human Development Report 2025. United Nations Development Programme. https://hdr.undp.org/content/human-development-report-2025

UNESCO (2023). Guidance for Generative AI in Education and Research. https://unesdoc.unesco.org/ark:/48223/pf0000386693

United Nations (2025). The Sustainable Development Goals Report 2025. United Nations, Department of Economic and Social Affairs. https://unstats.un.org/sdgs/report/2025/

WatchGuard (2025). Evasive Malware Surges 40 % in WatchGuard's Latest Internet Security Report. https://www.watchguard.com/es/wgrd-news/blog/evasive-malware-surges-40-watchguards-latest-internet-security-report

WIPO (2025). WIPO Conversation on Intellectual Property and Frontier Technologies. https://www.wipo.int/en/web/frontier-technologies/frontier_conversation

World Bank (2024). Global Trends in AI Governance. https://documents1.worldbank.org/curated/en/099120224205026271/pdf/P1786161ad76ca0ae1ba3b1558ca4ff88ba.pdf

World Bank (2025). World Development Report 2025: Standards for Development. World Bank. https://www.worldbank.org/en/publication/wdr2025

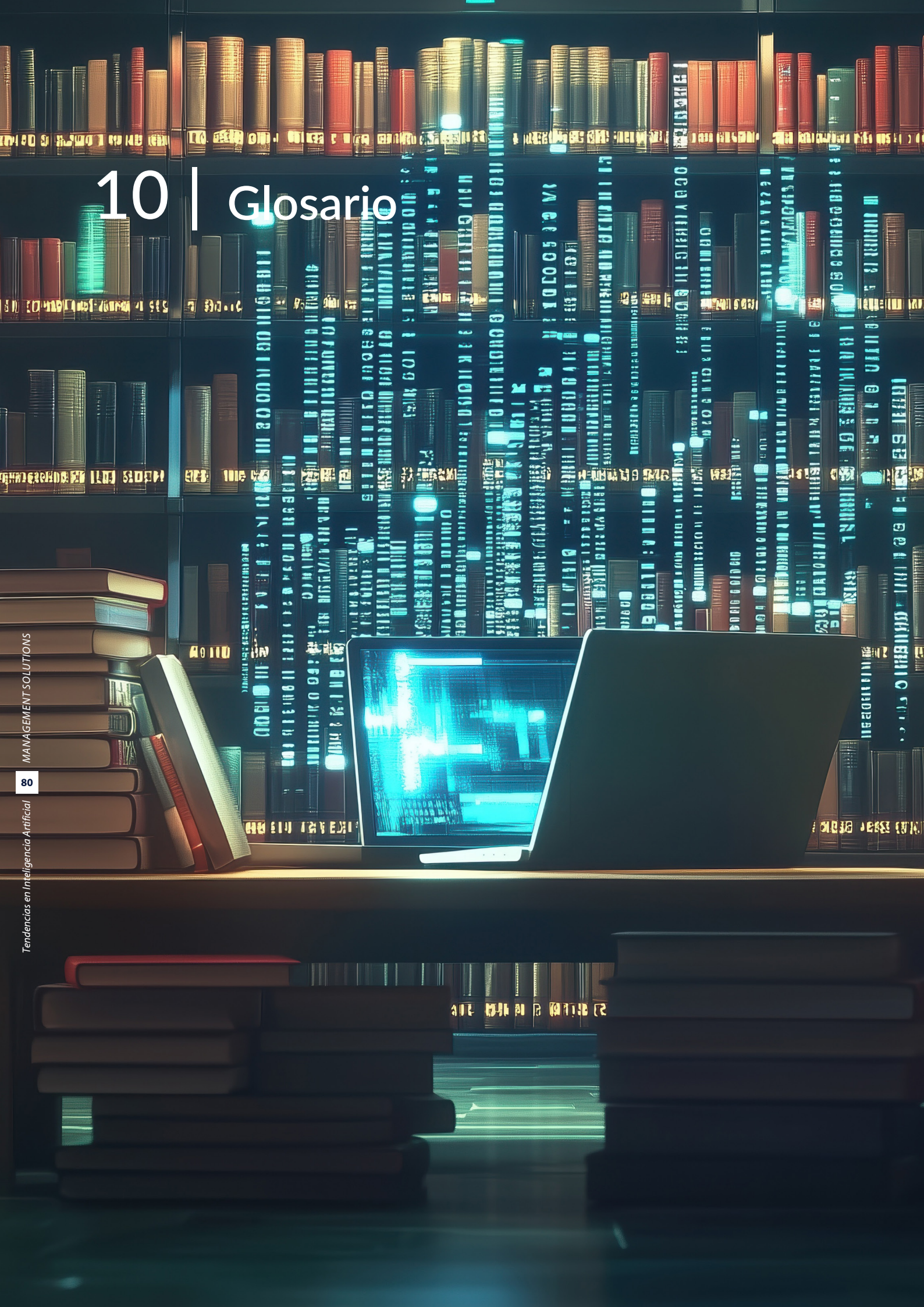
World Economic Forum (2025a). AI in Action: Beyond Experimentation to Transform Industry. https://reports.weforum.org/docs/WEF_AI_in_Action_Beyond_Experimentation_to_Transform_Industry_2025.pdf

World Economic Forum (2025b). Transforming Consumer Industries in the Age of AI. https://reports.weforum.org/docs/WEF_Transforming_Consumer_Industries_in_the_Age_of_AI_2025.pdf

World Health Organization (2024). Ethics and Governance of Artificial Intelligence for Health. https://iris.who.int/server/api/core/bitstreams/f780d926-4ae3-42ce-a6d6-e898a5562621/content

Yuan et al. (2025). The Impact of AI Adoption in the Workplace on Employees: A Systematic Review. https://www.researchgate.net/publication/396219396_The_Impact_of_AI_Adoption_in_the_Workplace_on_Employees_A_Systematic_Review

10 | Glosario



AGI (Artificial General Intelligence): Sistema de IA capaz de realizar cualquier tarea cognitiva que pueda realizar un ser humano, con generalización transferible entre dominios. Horizonte estratégico debatido cuya definición precisa carece de consenso científico.

AI Act: Reglamento (UE) 2024/1689, primer marco legal integral sobre IA. Clasifica los sistemas por nivel de riesgo (inaceptable, alto, limitado, mínimo) e impone obligaciones estructurales a los de alto riesgo. Sanciones de hasta 35 M€ o el 7 % de la facturación global.

AI Board: Foro de coordinación e interpretación común entre la Comisión Europea y las autoridades nacionales de supervisión de IA, creado por el AI Act.

AI Office: Órgano técnico central de la Comisión Europea responsable de la supervisión de modelos de IA de propósito general (GPAI) en el marco del AI Act.

AI-enhanced: Modelo organizativo en el que la IA se usa para optimizar procesos existentes sin rediseñarlos desde las capacidades de la IA. Estadio predominante en la mayoría de organizaciones actuales.

AI-first: Modelo organizativo en el que los procesos y la estructura se diseñan partiendo de las capacidades de la IA, asignando al juicio humano únicamente las tareas donde su ventaja comparativa es inequívoca.

AI-only: Modelo organizativo hipotético en el que las operaciones centrales prescinden funcionalmente del trabajo humano.

AIMS (AI Management System): Sistema de gestión de IA conforme a ISO/IEC 42001, equivalente a ISO 27001 para ciberseguridad pero específico para IA: cubre políticas, evaluaciones de impacto, control de proveedores y supervisión continua.

Alucinación: Comportamiento intrínseco de los modelos generativos por el que producen información falsa o inexacta presentada con aparente confianza. No es un bug ocasional, sino una consecuencia estructural del diseño actual de los LLMs.

Ambient AI (IA ambiental): IA que opera sin ser invocada explícitamente: observa el contexto de forma continua, infiere necesidades y actúa de forma proactiva. La interfaz desaparece; el entorno mismo se convierte en el punto de interacción.

Ambient scribe: Sistema de Ambient AI desplegado en entornos clínicos que escucha la conversación médico-paciente y genera automáticamente la documentación del encuentro sin instrucción explícita del usuario.

BEC (Business Email Compromise): Tipo de ciberataque en el que se suplanta la identidad de un directivo o proveedor para desviar fondos o extraer información. La IA generativa lo potencia mediante deepfakes de audio y vídeo de alta credibilidad.

Brecha de absorción: Distancia creciente entre la curva de capacidad tecnológica de la IA (exponencial, autoacelerada) y la curva de absorción organizativa (rediseño de procesos, reconversión de roles, gobernanza).

CAIO / CDAIO (Chief AI Officer / Chief Data and AI Officer): Función ejecutiva responsable del liderazgo estratégico de la IA en una organización.

Citizen data scientist: Profesional no técnico capaz de realizar análisis de datos básicos con herramientas visuales. La IA generativa supera este concepto al entregar capacidades analíticas avanzadas directamente a usuarios finales sin formación técnica.

Computación cuántica: Paradigma computacional que explota propiedades cuánticas (superposición, entrelazamiento) para resolver ciertos problemas intratables para sistemas clásicos.

Criptografía resistente a ataques cuánticos (PQC): Conjunto de algoritmos criptográficos diseñados para resistir ataques de ordenadores cuánticos. El NIST publicó los primeros estándares en 2024 (FIPS 203, 204, 205).

Dark LLM: Modelos de lenguaje modificados específicamente para cibercrimen (WormGPT, FraudGPT, GhostGPT). Generan malware, exploits y campañas de ingeniería social sin restricciones éticas. Comercializados en dark web con soporte técnico.

Data poisoning: Ataque adversario que consiste en inyectar datos maliciosos en el conjunto de entrenamiento de un modelo para degradar su comportamiento o introducir sesgos controlados por el atacante.

Deepfake: Contenido audiovisual sintético generado por IA que suplanta la apariencia o voz de una persona real. Usado en ataques BEC, fraude y desinformación con tasa de éxito significativamente superior al phishing tradicional.

Differential privacy: Técnica que añade ruido estadístico controlado a los datos o resultados para impedir la identificación de individuos, preservando utilidad estadística agregada.

DPIA (Data Protection Impact Assessment): Evaluación de impacto en protección de datos exigida por GDPR (Art. 35). El EDPB la considera obligatoria en la mayoría de despliegues de LLMs dado su procesamiento sistémico de datos personales.

Edge AI / TinyML: Capacidad de ejecutar modelos de IA directamente en dispositivos como smartphones, wearables y sensores sin depender de conectividad con la nube. Habilitador clave de la Ambient AI.

Efecto Bruselas: Fenómeno por el que la regulación de la UE (AI Act, GDPR) obliga a empresas globales a adaptar sus productos a los estándares europeos por el volumen del mercado, exportando de facto esa regulación al resto del mundo.

Ethics washing: Fenómeno por el que una organización publica principios éticos de IA sin traducirlos en controles operativos, responsabilidades concretas o mecanismos de auditoría.

Evasión adversaria: Ataque que introduce perturbaciones imperceptibles en los inputs de un modelo para provocar clasificaciones o respuestas incorrectas durante la inferencia, sin modificar el modelo en sí.

Explicabilidad: Capacidad de describir el funcionamiento interno de un modelo de IA de forma comprensible para diferentes audiencias (regulador, cliente, empleado). Requisito técnico en modelos regulados; implementable en ML con técnicas como SHAP, LIME o análisis de sensibilidad.

Feature engineering: Proceso de construir variables predictivas a partir de datos crudos para alimentar modelos de ML. Una de las fases más intensivas en conocimiento experto del ciclo de vida del ML clásico; acelerada significativamente por IA generativa.

Federated learning: Paradigma de entrenamiento distribuido en el que los datos permanecen en los dispositivos locales y solo se comparten actualizaciones del modelo. Mitiga riesgos de privacidad en LLMs a costa de mayor complejidad operativa.

GDPR (General Data Protection Regulation): Reglamento (UE) 2016/679 de protección de datos. Sus principios de minimización, derecho al olvido y transparencia presentan tensiones estructurales con la arquitectura y el ciclo de vida de los LLMs.

Gemelo digital (Digital Twin): Simulación dinámica de un sistema físico o humano que se actualiza en tiempo real con datos del sistema real. Funciona con alta fiabilidad en sistemas físicos deterministas; los LLMs han abierto su extensión a comportamiento humano y sistemas sociales complejos.

GPAI (General Purpose AI): Modelo de IA de propósito general (e.g., GPT, Claude, Gemini) capaz de realizar una amplia variedad de tareas.

Gradient boosting: Técnica de ML que combina múltiples árboles de decisión de forma secuencial para mejorar la predicción.

Harvest now, decrypt later: Estrategia por la que actores estatales capturan hoy comunicaciones cifradas con la intención de descifrarlas cuando la computación cuántica madure. Amenaza presente para datos con larga vida útil de confidencialidad.

Hub & spokes: Modelo organizativo de IA con un Centro de Excelencia central que establece capacidades transversales, mientras equipos descentralizados en líneas de negocio desarrollan soluciones específicas con reporte funcional cruzado al hub.

IA agéntica: Sistemas de IA que planifican, ejecutan tareas complejas y operan de forma autónoma sobre infraestructuras corporativas reales, más allá de responder a prompts. Capacidades incrementales: estado persistente, planificación dinámica, ejecución sobre sistemas reales y orquestación multiagente.

IA generativa: Familia de modelos capaces de generar contenido original (texto, imágenes, audio, vídeo, código) ante instrucciones en lenguaje natural. Basada en arquitecturas transformer de gran escala.

IRB (Internal Ratings-Based): Enfoque regulatorio que permite a las entidades bancarias usar modelos internos para calcular el capital regulatorio por riesgo de crédito.

Jagged intelligence (inteligencia dentada): Concepto de Andrej Karpathy para describir el perfil de capacidades de los LLMs actuales: brillantes en tareas complejas (olimpiadas matemáticas) y frágiles en tareas aparentemente simples (comparar números decimales).

Jailbreak: Técnica para eludir los controles de seguridad de un modelo de IA mediante instrucciones diseñadas para hacer que el sistema ignore sus restricciones de comportamiento.

LIME (Local Interpretable Model-agnostic Explanations): Técnica de explicabilidad que genera aproximaciones locales de un modelo complejo para explicar predicciones individuales.

LLM (Large Language Model): Modelo de lenguaje de gran escala basado en arquitectura transformer, entrenado sobre vastos corpus textuales. Fundamento técnico de parte de la IA generativa actual. Ejemplos: GPT, Claude, Gemini, Llama.

LLMOps (Large Language Model Operations): Extensión de MLOps específica para gestionar las propiedades de los LLMs en producción: comportamiento no determinista, prompts como superficie de riesgo, alucinaciones, costes por token y trazabilidad de interacciones.

Lock-in de proveedor: Dependencia estructural de un único proveedor de modelos o infraestructura que dificulta la migración (reescritura de integraciones, recertificación de compliance, costes prohibitivos). Riesgo estratégico en la adopción de IA.

LoRA (Low-Rank Adaptation): Técnica eficiente de fine-tuning de LLMs que adapta el modelo base a un dominio específico modificando solo un subconjunto de parámetros, reduciendo drásticamente los recursos computacionales necesarios.

Machine learning (ML): Rama de la IA que desarrolla algoritmos capaces de aprender patrones a partir de datos históricos sin ser programados explícitamente para cada tarea. A diferencia de la IA generativa, los modelos de ML clásico clasifican, predicen y optimizan; no generan contenido.

Machine unlearning: Conjunto de técnicas experimentales que buscan eliminar selectivamente el conocimiento derivado de datos específicos de un modelo ya entrenado, en respuesta al derecho al olvido del GDPR.

Malware polimórfico: Código malicioso que muta su firma para evadir detección. La IA generativa lo potencia: variantes avanzadas reescriben su código cada 15 segundos manteniendo funcionalidad idéntica, y vencen a sistemas basados en firmas estáticas.

Many hands problem: Problema de responsabilidad difusa en sistemas complejos donde el daño se produce sin intención deliberada y la cadena causal se fragmenta entre múltiples actores (diseñadores, entrenadores, desplegadores, usuarios).

MCP (Model Context Protocol): Estándar abierto que normaliza cómo los modelos de IA interactúan con aplicaciones, fuentes de datos y herramientas externas. Elimina la deuda técnica de las integraciones propietarias y convierte cada herramienta en un activo reutilizable por cualquier agente.

MLOps (Machine Learning Operations): Conjunto de procesos estandarizados y capacidades tecnológicas para construir, desplegar y operacionalizar modelos de ML de forma fiable a lo largo de todo su ciclo de vida. Cubre preparación de datos, experimentación, validación, despliegue y monitorización.

Model drift: Degradación silenciosa del rendimiento de un modelo en producción causada por cambios en la distribución de los datos de entrada respecto a los de entrenamiento.

Multimodalidad: Capacidad de un modelo de IA de procesar y generar simultáneamente múltiples tipos de información (texto, imágenes, audio, vídeo, código) en una sola arquitectura conversacional.

NIST AI RMF: Marco de gestión de riesgos de IA del National Institute of Standards and Technology. Organizado en las funciones Govern-Map-Measure-Manage. Orientado a características de IA fiable (trustworthy AI).

Orquestación multiagente: Arquitectura en la que múltiples agentes de IA especializados colaboran bajo un coordinador central para alcanzar objetivos complejos, gestionando dependencias, prioridades y traspaso de información entre agentes.

Phishing hiperpersonalizado: Ataques de phishing generados por IA que analizan perfiles públicos y estilo de escritura corporativa para crear mensajes altamente personalizados.

Prompt: Instrucción o entrada en lenguaje natural que un usuario proporciona a un sistema de IA generativa para obtener una respuesta.

Prompt injection: Ataque en el que instrucciones maliciosas embebidas en los inputs (documentos, URLs, datos externos) manipulan el comportamiento del modelo para ignorar restricciones o ejecutar acciones no autorizadas.

RAG (Retrieval-Augmented Generation): Arquitectura que complementa un LLM con un sistema de recuperación de información externa en tiempo real. Reduce alucinaciones al anclar las respuestas en documentos verificables, y separa el conocimiento actualizable de la memoria estática del modelo.

ReAct (Reason + Act): Bucle de control de agentes de IA que combina razonamiento (análisis de la situación) y acción (uso de herramientas o APIs) de forma iterativa. Fundamento del núcleo cognitivo de los sistemas agénticos.

Reskilling: Proceso de adquisición de nuevas competencias para desempeñar roles profesionales diferentes a los actuales, en respuesta a la transformación del trabajo por la IA. Palanca estratégica dado el desequilibrio estructural entre oferta y demanda de talento en IA.

Robótica humanoide: Robots con forma y capacidades de movimiento similares a las humanas, integrados con modelos de IA para percepción, razonamiento y aprendizaje.

Shadow AI: Uso no autorizado de herramientas de IA generativa en entornos corporativos, con frecuencia en plataformas públicas sin garantías de privacidad.

SHAP (SHapley Additive exPlanations): Técnica de explicabilidad basada en teoría de juegos que cuantifica la contribución de cada variable a una predicción individual de un modelo.

SIEM / XDR / SOAR: Plataformas de ciberseguridad: SIEM (Security Information and Event Management), XDR (Extended Detection and Response) y SOAR (Security Orchestration, Automation and Response).

Soberanía tecnológica: Capacidad de un estado u organización de controlar las capas críticas de la cadena de valor de la IA (hardware, infraestructura, modelos, talento) para mantener autonomía estratégica.

Tecnobloques: Esferas de influencia tecnológica parcialmente incompatibles (lideradas por EEUU, China y Europa) con lógicas propias de gobernanza, seguridad y valores.

Test de Turing: Criterio propuesto por Alan Turing (1950) para evaluar si una máquina exhibe comportamiento inteligente indistinguible del humano en conversación.

Transformer: Arquitectura de red neuronal escalable basada en mecanismos de atención, introducida por Google en 2017. Fundamento técnico de todos los LLMs modernos.

UEBA (User and Entity Behavior Analytics): Sistemas de ciberseguridad que establecen líneas base dinámicas de comportamiento normal y detectan anomalías.

Upskilling: Proceso de ampliación de competencias existentes para adaptarse a nuevos requisitos del mismo rol profesional, específicamente en el contexto de la adopción de IA.

Vibe coding: Paradigma de desarrollo de software por conversación iterativa en lenguaje natural con sistemas de IA que interpretan requisitos, generan aplicaciones completas, detectan errores y producen tests y documentación de forma automática. Término acuñado por Andrej Karpathy.