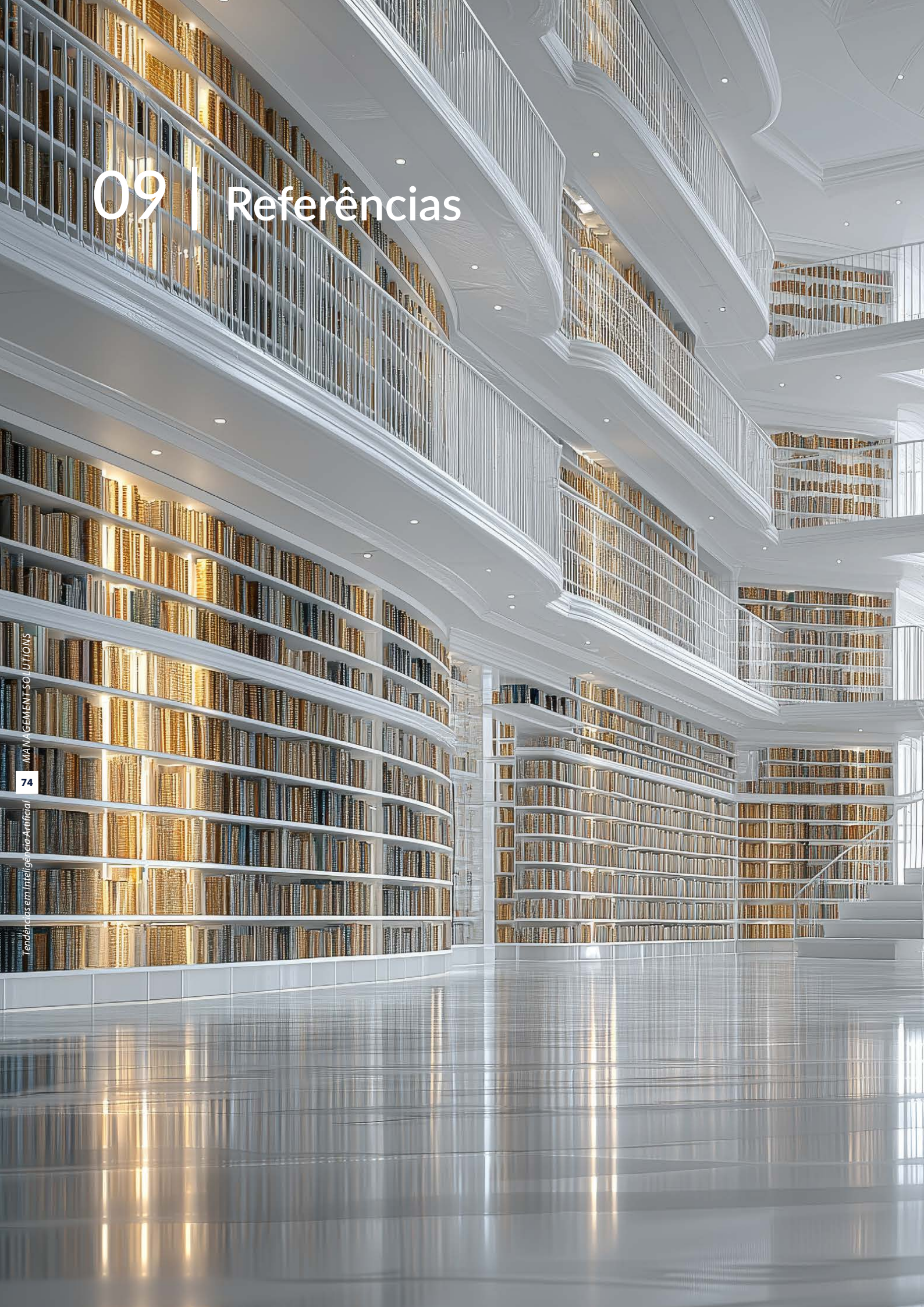
09 | Referências



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 | Glossário

AGI (Artificial General Intelligence): um sistema de IA capaz de realizar qualquer tarefa cognitiva que possa ser realizada por um ser humano, com generalização transferível entre domínios. Um horizonte estratégico debatido cuja definição precisa carece de consenso científico.

AI Act: Regulamento (UE) 2024/1689, a primeira estrutura jurídica abrangente sobre IA. Classifica os sistemas por nível de risco (inaceitável, alto, limitado, limitado, mínimo) e impõe obrigações estruturais aos sistemas de alto risco. Penalidades de até €35 milhões ou 7% do faturamento global.

AI Board: Fórum para coordenação e interpretação comum entre a Comissão Europeia e as autoridades nacionais de supervisão de IA, criado pelo AI Act.

AI Office: órgão técnico central da Comissão Europeia responsável pela supervisão de modelos de IA de uso geral (GPAI) nos termos do AI Act.

AI-enhanced: modelo organizacional no qual a IA é usada para otimizar os processos existentes sem redesenhá-los a partir de recursos de IA. Estágio predominante na maioria das organizações atualmente.

AI-first: modelo organizacional no qual os processos e a estrutura são projetados com base nos recursos de IA, atribuindo o julgamento humano apenas a tarefas em que sua vantagem comparativa é inequívoca.

AI-only: um modelo organizacional hipotético no qual as operações principais dispensam funcionalmente o trabalho humano.

AIMS (AI Management System): sistema de gerenciamento de IA de acordo com a ISO/IEC 42001, equivalente à ISO 27001 para segurança cibernética, mas específico para IA: abrange políticas, avaliações de impacto, controle de fornecedores e monitoramento contínuo.

Alucinação: comportamento intrínseco de modelos generativos por meio do qual eles produzem informações falsas ou imprecisas apresentadas com aparente confiança. Não se trata de um bug ocasional, mas de uma consequência estrutural do projeto atual dos LLMs.

Ambient AI: IA que opera sem ser explicitamente invocada: observa continuamente o contexto, infere necessidades e age proativamente. A interface desaparece; o próprio ambiente se torna o ponto de interação.

Ambient scribe: Sistema de IA ambiente implantado em ambientes clínicos que ouve a conversa entre o médico e o paciente e gera automaticamente a documentação do encontro sem instruções explícitas do usuário.

BEC (Business Email Compromise): um tipo de ataque cibernético no qual a identidade de um gerente ou fornecedor é personificada para desviar fundos ou extrair informações. A IA generativa possibilita isso por meio de *deepfakes* de áudio e vídeo altamente confiáveis.

Lacuna de absorção: distância cada vez maior entre a curva de capacidade da tecnologia de IA (exponencial, autoacelerada) e a curva de absorção organizacional (redesenho de processos, reengenharia de funções, governança).

CAIO / CDAIO (Chief AI Officer / Chief Data and AI Officer): função executiva responsável pela liderança estratégica da IA em uma organização.

Citizen data scientist: profissional não técnico capaz de realizar análises básicas de dados com ferramentas visuais. A IA generativa vai além desse conceito, fornecendo recursos analíticos avançados diretamente para usuários finais não técnicos.

Computação quântica: paradigma computacional que explora as propriedades quânticas (superposição, emaranhamento) para resolver determinados problemas intratáveis para sistemas clássicos.

Criptografia resistente a ataques quânticos (PQC): conjunto de algoritmos criptográficos projetados para resistir a ataques de computadores quânticos. O NIST publicou os primeiros padrões em 2024 (FIPS 203, 204, 205).

Dark LLM: modelos de linguagem modificados especificamente para o crime cibernético (WormGPT, FraudGPT, GhostGPT). Eles geram malware, explorações e campanhas de engenharia social sem restrições éticas. Comercializados na dark web com suporte técnico.

Data poisoning: ataque adversário que consiste em injetar dados maliciosos no conjunto de treinamento de um modelo para degradar seu comportamento ou introduzir vieses controlados pelo invasor.

Deepfake: conteúdo audiovisual sintético gerado por IA que imita a aparência ou a voz de uma pessoa real. Usado em ataques de BEC, fraude e desinformação com uma taxa de sucesso significativamente maior do que o phishing tradicional.

Differential privacy: técnica que adiciona ruído estatístico controlado aos dados ou resultados para impedir a identificação de indivíduos, preservando a utilidade estatística agregada.

DPIA (Data Protection Impact Assessment): avaliação do impacto da proteção de dados exigida pelo GDPR (Art. 35). O EDPB a considera obrigatória na maioria das implantações de LLM, dado seu processamento sistêmico de dados pessoais.

Edge AI / TinyML: capacidade de executar modelos de IA diretamente em dispositivos como smartphones, wearables e sensores sem depender de conectividade com a nuvem. Principal facilitador da IA ambiente.

Efeito Bruxelas: fenômeno pelo qual a regulação da UE (AI Act, GDPR) força as empresas globais a adaptarem seus produtos aos padrões europeus devido ao volume de mercado, exportando de fato essa regulação para o resto do mundo.

Ethics washing: fenômeno em que uma organização publica princípios éticos de IA sem traduzi-los em controles operacionais, responsabilidades concretas ou mecanismos de auditoria.

Evasão adversária: um ataque que introduz perturbações imperceptíveis nas entradas de um modelo para causar classificações ou respostas incorretas durante a inferência, sem modificar o próprio modelo.

Explicabilidade: capacidade de descrever o funcionamento interno de um modelo de IA de forma que seja compreensível para diferentes públicos (regulador, cliente, funcionário). Requisito técnico em modelos regulados; implementável em ML com técnicas como SHAP, LIME ou análise de sensibilidade.

Feature engineering: processo de criação de variáveis preditivas a partir de dados brutos para alimentar modelos de AM. Uma das fases mais intensivas em conhecimento especializado do ciclo de vida clássico de AM; significativamente acelerada pela IA generativa.

Federated learning: paradigma de treinamento distribuído no qual os dados permanecem em dispositivos locais e somente as atualizações de modelos são compartilhadas. Atenua os riscos de privacidade em LLMs ao custo de maior complexidade operacional.

GDPR (General Data Protection Regulation): Regulamento de Proteção de Dados (UE) 2016/679. Seus princípios de minimização, direito de ser esquecido e transparência apresentam tensões estruturais com a arquitetura e o ciclo de vida dos LLMs.

Gêmeo Digital (Digital Twin): uma simulação dinâmica de um sistema físico ou humano que é atualizada em tempo real com dados do sistema real. Funciona com alta confiabilidade em sistemas físicos determinísticos; os LLMs abriram sua extensão para o comportamento humano e sistemas sociais complexos.

GPAI (General Purpose AI): modelo de IA de propósito geral (por exemplo, GPT, Claude, Gemini) capaz de executar uma ampla variedade de tarefas.

Gradient boosting: técnica de ML que combina várias árvores de decisão sequencialmente para melhorar a previsão.

Harvest now, decrypt later: estratégia em que agentes estatais capturam comunicações criptografadas hoje com a intenção de descriptografá-las quando a computação quântica amadurecer. Ameaça presente aos dados com uma longa vida útil de confidencialidade.

Hub & spokes: modelo organizacional de IA com um Centro de Excelência central que estabelece recursos transversais, enquanto equipes descentralizadas em linhas de negócios desenvolvem soluções específicas com relatórios multifuncionais para o hub.

IA agêntica: sistemas de IA que planejam, executam tarefas complexas e operam de forma autônoma em infraestruturas corporativas reais, além de responder a avisos. Recursos incrementais: estado persistente, planejamento dinâmico, execução em sistemas reais e orquestração de vários agentes.

IA generativa: família de modelos capazes de gerar conteúdo original (texto, imagens, áudio, vídeo, código) em resposta a instruções de linguagem natural. Baseada em arquiteturas de transformadores em larga escala.

IRB (Internal Ratings-Based): abordagem regulatória que permite que os bancos usem modelos internos para calcular o capital regulatório para o risco de crédito.

Jagged intelligence: conceito de Andrej Karpathy para descrever o perfil de habilidades dos LLMs atuais: brilhantes em tarefas complexas (olimpíadas de matemática) e frágeis em tarefas aparentemente simples (comparação de números decimais).

Jailbreak: uma técnica para contornar os controles de segurança de um modelo de IA por meio de instruções projetadas para fazer com que o sistema ignore suas restrições comportamentais.

LIME (Local Interpretable Model-agnostic Explanations): técnica de explicabilidade que gera aproximações locais de um modelo complexo para explicar previsões individuais.

LLM (Large Language Model): modelo de linguagem em grande escala baseado na arquitetura do transformador, treinado em vastos corpora textuais. Base técnica de parte da IA generativa atual. Exemplos: GPT, Claude, Gemini, Llama.

LLMOps (Large Language Model Operations): extensão específica do MLOps para gerenciar as propriedades dos LLMs em produção: comportamento não determinístico, avisos como superfície de risco, alucinações, custo por token e rastreabilidade das interações.

Lock-in de fornecedor: dependência estrutural de um único modelo ou fornecedor de infraestrutura que dificulta a migração (reescrever integrações, recertificação de compliance, custos proibitivos). Risco estratégico na adoção de IA.

LoRA (Low-Rank Adaptation): técnica eficiente de ajuste fino de LLM que adapta o modelo básico a um domínio específico modificando apenas um subconjunto de parâmetros, reduzindo drasticamente os recursos computacionais necessários.

Machine learning (ML): ramo da IA que desenvolve algoritmos capazes de aprender padrões a partir de dados históricos sem serem explicitamente programados para cada tarefa. Diferentemente da IA generativa, os modelos clássicos de ML classificam, preveem e otimizam; eles não geram conteúdo.

Machine unlearning: conjunto de técnicas experimentais que buscam remover seletivamente o conhecimento derivado de dados específicos de um modelo já treinado, em resposta ao direito do GDPR de ser esquecido.

Malware polimórfico: código malicioso que altera sua assinatura para evitar a detecção. A IA generativa é responsável por isso: variantes avançadas reescrevem seu código a cada 15 segundos, mantendo a funcionalidade idêntica, e derrotam sistemas baseados em assinaturas estáticas.

Many hands problem: problema de responsabilidade difusa em sistemas complexos em que o dano ocorre sem intenção deliberada e a cadeia causal é fragmentada entre vários atores (designers, treinadores, implantadores, usuários).

MCP (Model Context Protocol): padrão aberto que padroniza como os modelos de IA interagem com aplicativos, fontes de dados e ferramentas externas. Ele elimina o débito técnico de integrações proprietárias e torna cada ferramenta um ativo reutilizável para qualquer agente.

MLOps (*Machine Learning* **Operations)**: um conjunto de processos padronizados e recursos tecnológicos para criar, implantar e operacionalizar modelos de ML de forma confiável durante todo o seu ciclo de vida. Ele abrange a preparação de dados, a experimentação, a validação, a implementação e o monitoramento.

Model drift: degradação silenciosa do desempenho de um modelo de produção causada por alterações na distribuição dos dados de entrada em relação aos dados de treinamento.

Multimodalidade: a capacidade de um modelo de IA de processar e gerar simultaneamente vários tipos de informações (texto, imagens, áudio, vídeo, código) em uma única arquitetura de conversação.

NIST AI RMF: estrutura de gerenciamento de riscos de IA do Instituto Nacional de Padrões e Tecnologia. Organizada nas funções

Governar, Mapear, Medir e Gerenciar. Focado em recursos de IA confiáveis.

Orquestração multiagente: Arquitetura na qual vários agentes de IA especializados colaboram com um coordenador central para atingir objetivos complexos, gerenciando dependências, prioridades e transferências de informações entre agentes.

Phishing hiperpersonalizado: ataques de *phishing* gerados por IA que analisam perfis públicos e estilo de redação corporativa para criar mensagens altamente personalizadas.

Prompt: uma instrução ou entrada de linguagem natural que um usuário fornece a um sistema de IA generativo para obter uma resposta.

Prompt injection: ataque no qual instruções maliciosas incorporadas em entradas (documentos, URLs, dados externos) manipulam o comportamento do modelo para ignorar restrições ou executar ações não autorizadas.

RAG (Retrieval-Augmented Generation): arquitetura que complementa um LLM com um sistema de recuperação de informações externas em tempo real. Ela reduz as alucinações ancorando as respostas em documentos verificáveis e separa o conhecimento atualizável da memória estática do modelo.

ReAct (Reason + Act): loop de controle de agente de IA que combina raciocínio (análise de situação) e ação (uso de ferramentas ou APIs) iterativamente. A base do núcleo cognitivo dos sistemas agênticos.

Reskilling: o processo de aquisição de novas competências para desempenhar funções profissionais diferentes das atuais, em resposta à transformação do trabalho pela IA. Alavanca estratégica devido ao desequilíbrio estrutural entre a oferta e a demanda de talentos em IA.

Robótica humanoide: robôs com forma e capacidade de movimento semelhantes aos humanos, integrados a modelos de IA para percepção, raciocínio e aprendizado.

Shadow AI: uso não autorizado de ferramentas de IA generativas em ambientes corporativos, geralmente em plataformas públicas sem garantias de privacidade.

SHAP (SHapley Additive exPlanations): técnica de explicabilidade baseada na teoria dos jogos que quantifica a contribuição de cada variável para uma previsão individual de um modelo.

SIEM / XDR / SOAR: plataformas de segurança cibernética: SIEM (Security Information and Event Management), XDR (Extended Detection and Response) e SOAR (Security Orchestration, Automation and Response).

Soberania tecnológica: a capacidade de um estado ou organização de controlar camadas essenciais da cadeia de valor da IA (hardware, infraestrutura, modelos, talentos) para manter a autonomia estratégica.

Technoblocks: esferas parcialmente incompatíveis de influência tecnológica (lideradas pelos EUA, China e Europa) com suas próprias lógicas de governança, segurança e valores.

Teste de Turing: Critério proposto por Alan Turing (1950) para avaliar se uma máquina apresenta comportamento inteligente indistinguível da conversação humana.

Transformer: arquitetura de rede neural escalável baseada em mecanismos de atenção, introduzida pelo Google em 2017. Base técnica de todos os LLMs modernos.

UEBA (User and Entity Behaviour Analytics): sistemas de segurança cibernética que estabelecem linhas de base dinâmicas de comportamento normal e detectam anomalias.

Upskilling: o processo de expansão das competências existentes para se adaptar aos novos requisitos da mesma função profissional, especificamente no contexto da adoção de IA.

Vibe coding: um paradigma de desenvolvimento de software por meio de conversas iterativas em linguagem natural com sistemas de IA que interpretam requisitos, geram aplicativos completos, detectam erros e produzem testes e documentação automaticamente. Termo cunhado por Andrej Karpathy.