

AI-AUGMENTED PROJECT MANAGEMENT OFFICES: A SCOPING REVIEW OF DELIVERABLE AUTOMATION IN PROFESSIONAL SERVICES CONSULTING

Cesar Lopes Maçol Costa¹

Abstract: Professional services firms face growing margin pressure. Clients want more deliverables for smaller budgets, which forces firms to rethink how the Project Management Office (PMO) produces status reports, meeting minutes, schedules, and first drafts of technical documents. Generative artificial intelligence (AI) is emerging as a possible structural response to this pressure. This paper presents a scoping review of the academic and industry literature on automating PMO deliverables with generative AI, guided by the transparency principles of PRISMA and its scoping review extension (PRISMA-ScR). The review maps the deliverable types that prior work documents as automatable and the productivity gains it reports, which range from about 12% to 40% depending on the task. It sets these gains against evidence of quality risk and overreliance on automation. The findings are framed as theoretical propositions that require empirical validation. They show a fragmented literature and no framework tested in a real multi-client consulting environment. The paper closes with a research agenda on quality governance for generative AI applied to corporate deliverables. Because a single reviewer screened open-access sources, the review does not claim exhaustive coverage.

Keywords: Generative artificial intelligence. Project management office. Deliverable automation. Consulting. Scoping review.

¹ Master of Science in Business Analytics (MSBA) — Indiana University, Kelley School of Business | Ago/2023 | GPA: 4.000/4.000 Master of Business Administration (MBA) — Indiana University, Kelley School of Business | Mai/2022 | GPA: 3.992/4.000 | Academic Excellence Achievement — Highest Distinction Bacharelado em Engenharia de Produção — Universidade Federal Fluminense (UFF) | Ago/2006 | GPA: 8,04/10,00

INTRODUCTION

Context and research problem

The consulting industry is repricing its services. The Everest Group (2025) market report says the sector entered 2025 on a gradual and uneven recovery after a significant slowdown in 2024. Firms are moving away from hourly billing toward asset-based and value-led models, and they are restructuring delivery to meet buyers' demands for speed, scale, and sector expertise. The pressure on margins is real. Clients expect a growing volume of deliverables, such as status reports, meeting minutes, schedules, analyses, and first drafts of technical documents, without a matching increase in budget.

The traditional PMO consolidates these deliverables and relies heavily on repetitive manual work. That reliance makes it a bottleneck as demand for scale grows. Generative AI is a candidate for a structural fix, not just a personal productivity aid. Industry surveys point the same way. In its first-quarter 2024 survey of 2,835 executives, the Deloitte AI Institute (2024a) found that generative AI efforts focused on efficiency, productivity, and cost reduction more than on innovation. By the third quarter, more than half of the 2,770 respondents also cited benefits beyond efficiency, and 55% reported avoiding some use cases because of data problems (DELOITTE AI INSTITUTE, 2024b). The Project Management Institute (PMI, 2024) surveyed 500 project professionals in 18 sectors and 12 countries. It reports that nearly two in five now use generative AI in more than half of their project work, and that this use grew 86% since its previous survey. These sources are self-reported industry surveys without control groups. This paper treats them as descriptive evidence of a market trend, not as proof of causal effects (see Section 3.2.1).

Research gap

Despite this industry evidence, the academic literature on generative AI in consulting

remains generic and fragmented. Many studies examine professional writing in general (NOY; ZHANG, 2023) or a single domain such as customer support (BRYNJOLFSSON; LI; RAYMOND, 2025). Neither connects its findings to the PMO deliverable lifecycle. The study closest to consulting is the field experiment with Boston Consulting Group consultants by Dell’Acqua et al. (2026), but it evaluates consulting tasks broadly and does not separate the deliverables a PMO typically produces. Few studies synthesize productivity evidence from multi-client settings. Most of what is publicly available consists of anecdotal case studies or industry reports that do not state how evidence was selected.

Objective

This paper uses a scoping review of the available academic and technical literature to map three things: which stages of the PMO deliverable lifecycle have been documented as automatable with generative AI, with what productivity results, and under what conditions for success. It organizes the evidence as theoretical propositions that future studies can test in a real multi-client consulting environment. The review does not claim to be exhaustive. It did not search licensed indexed databases, and Sections 3.1, 3.9, and 5 discuss this limitation.

Research questions

RQ1: Which stages of the PMO deliverable lifecycle has the literature documented as automatable with generative AI?

RQ2: What productivity results (cycle time, delivery volume, rework reduction) do existing studies report?

RQ3: Which critical success factors and adoption barriers recur across studies?

Expected contribution

This paper offers a set of theoretical propositions, not yet a validated framework, for automating PMO deliverables. The propositions build on the author's earlier published work. Costa (2025) reports that applying generative AI to the Financial Planning & Analysis (FP&A) cycle was associated with a 45% reduction in budget-forecasting cycle time and a 38% improvement in forecast accuracy. This paper uses those figures only as an illustrative comparison point. They are not statistical evidence that the same effect will hold in the PMO domain. Because this is self-citation, the Conflict of Interest Statement at the end of the paper discloses it, and readers should weigh it when reading Section 4.1.

LITERATURE REVIEW

Digital transformation in professional services

Consulting is a knowledge-intensive, labor-intensive business. Its product is the skilled time of its professionals, packaged as deliverables. That makes pricing models, especially the choice between fixed fees and time-and-materials billing, very sensitive to productivity gains. When a task that used to take hours takes minutes, time-and-materials billing hurts the firm, and fixed fees require higher delivery volume to hold margins. The Everest Group (2025) report, which offers descriptive industry evidence rather than experimental evidence, describes this shift. Firms are moving toward asset-based and value-led models and restructuring delivery around speed, scale, and sector expertise.

Generative AI in corporate settings

Peer-reviewed experiments document large productivity gains that they attribute to generative AI itself. Noy and Zhang (2023) ran a randomized experiment, published in *Science*, in which college-educated professionals completed mid-level writing tasks. The participants included managers, HR

professionals, grant writers, marketers, consultants, and data analysts. Those with access to ChatGPT finished about 40% faster, and their work received quality ratings about 18% higher than the control group's. The gains were largest for initially lower-performing participants, a pattern known as skill leveling. Because the participants were online experiment subjects rather than practicing consultants, the results may not transfer directly to a PMO setting.

In customer support, Brynjolfsson, Li, and Raymond (2025) studied the staggered rollout of a generative conversational assistant to 5,172 agents at one Fortune 500 software company. Productivity, measured as issues resolved per hour, rose by about 15% on average. Less experienced and lower-skilled agents improved in both speed and quality, while the most experienced and highest-skilled agents gained slightly in speed and lost slightly in quality. Because the study covers a single company, it is unclear how far the results generalize. Industry evidence, which is descriptive rather than causal, adds context. The Deloitte AI Institute (2024a) found that most organizations still depend on off-the-shelf generative AI tools and that talent, governance, and risk management are the main gaps in organizational readiness.

Project management offices: function and evolution

The traditional PMO handles governance, control, and reporting. It standardizes processes, consolidates schedules, produces status reports, and enforces methodological compliance across projects. Industry data suggest that this role may be shifting from schedule control toward orchestrating automated deliverables, although the data are descriptive and self-reported. In a survey of 500 project professionals in 18 sectors and 12 countries, PMI (2024) found that nearly two in five now use generative AI in more than half of their project work, an 86% increase over its previous survey.

Human-in-the-loop oversight and quality of AI-generated output

The gains in Section 2.2 are neither uniform nor risk-free. Dell’Acqua et al. (2026) report a pre-registered field experiment with 758 Boston Consulting Group consultants, about 7% of the firm’s individual-contributor consulting workforce. The consultants worked on realistic management consulting tasks. The study introduces the “jagged technological frontier,” the uneven boundary between tasks where AI performs reliably and tasks where it does not, even when the tasks look similarly difficult. On 18 tasks inside the frontier, consultants using GPT-4 completed 12.2% more tasks, finished 25.1% faster, and delivered work of significantly higher quality. On a task deliberately placed outside the frontier, consultants using AI were 19 percentage points less likely to produce a correct solution.

Dell’Acqua (2022) documents a behavioral side of this risk in a pre-registered field experiment with 181 professional recruiters, each of whom reviewed 44 résumés. The recruiters received AI recommendations of randomly varied quality. Those who worked with a lower-quality AI were more accurate than those who worked with a higher-quality one. They spent more time and effort on each résumé and were less likely to accept the recommendation automatically. Dell’Acqua calls the opposite pattern “falling asleep at the wheel”: the better an AI appears, the less reason people have to stay critically engaged.

Doshi and Hauser (2024) ran an online experiment, published in *Science Advances*, in which some writers received story ideas from a large language model. Access to AI ideas made stories more creative, better written, and more enjoyable, especially for less creative writers. However, AI-assisted stories were more similar to one another than stories written by humans alone. In a multi-client PMO, this finding points to a risk that the consulting literature has not yet measured. Strategic deliverables for different clients could converge if firms do not calibrate their use of generative AI.

These behavioral risks sit on top of a technical one. Ji et al. (2023), in *ACM Computing Surveys*, review how text generation models produce output that is fluent but unsupported or unfaithful,

a problem commonly called hallucination. For a PMO that uses generative AI to draft deliverables for corporate clients, hallucination is a first-order risk. Automating PMO deliverables therefore requires human review calibrated to the task type and to its position relative to the frontier. That requirement leads to the governance framework proposed in Section 4.5.

Synthesis and positioning of the study

The literature reviewed supports three conclusions. First, experiments show real productivity gains from generative AI in knowledge work generally (NOY; ZHANG, 2023), in customer support (BRYNJOLFSSON; LI; RAYMOND, 2025), and in management consulting tasks (DELL'ACQUA et al., 2026). Second, experiments of equal weight document quality risks and overreliance (DELL'ACQUA, 2022; DOSHI; HAUSER, 2024), which rest on a known technical weakness (JI et al., 2023). Third, industry reports (PMI, 2024; DELOITTE AI INSTITUTE, 2024a, 2024b; EVEREST GROUP, 2025) indicate that these dynamics already operate at market scale, though they are descriptive and not causal. The literature still lacks a synthesis organized around the PMO deliverable lifecycle. It also lacks a replicable framework, tested in a real multi-client consulting environment, that joins productivity gains with quality control. This paper defines that gap.

METHODS

Study design

This paper is a scoping review, not a full systematic review. The distinction is deliberate. A full systematic review under PRISMA 2020 (PAGE et al., 2021) calls for an exhaustive search of several licensed indexed databases, independent double screening with a reported agreement statistic such as Cohen's kappa, a registered protocol, and a PRISMA flow diagram with counts for every stage. This study did not meet those conditions. A single reviewer did the screening. The search did

not cover Scopus or Web of Science, and no protocol was registered in advance. The study therefore follows the transparency and traceability principles of the PRISMA extension for scoping reviews, PRISMA-ScR (TRICCO et al., 2018), and uses a qualitative, descriptive approach with thematic synthesis. Section 4.5 proposes a full systematic review as the next step.

Search strategy

The search covered open-access academic sources (Google Scholar, SSRN, and arXiv) and reference publications from institutions and industry bodies (Harvard Business School working papers, the Project Management Institute, the Deloitte AI Institute, and the Everest Group). It combined the following terms in English and Portuguese: (“generative AI” OR “GenAI” OR “large language models”) AND (“project management office” OR “PMO”); (“AI automation” AND “professional services” AND “deliverable*”); (“digital transformation” AND “consulting” AND “artificial intelligence”); and (“human-in-the-loop” AND “generative AI” AND “quality”). Searches ran between September and October 2026 and covered publications from 2020 to 2026. A full registered review should extend this strategy to Scopus and Web of Science, which this exploratory phase did not consult.

Classification of evidence by strength

The sources differ widely in design, so this paper sorts the evidence into four tiers. Section 2 labels each source by tier where it matters, and Appendix A lists the sources one by one. Tier (i) is peer-reviewed or pre-registered experimental evidence: Noy and Zhang (2023), Brynjolfsson, Li, and Raymond (2025), Dell’Acqua et al. (2026), Dell’Acqua (2022), and Doshi and Hauser (2024). This tier carries the most causal weight, although Dell’Acqua (2022) is still a working paper. Tier (ii) is technical synthesis: Ji et al. (2023), which characterizes a structural risk but does not measure productivity. Tier (iii) is descriptive industry evidence, often called grey literature: PMI (2024), the

Deloitte AI Institute (2024a, 2024b), and the Everest Group (2025). These sources lack control groups, so the paper treats them as market context, not as causal proof. Tier (iv) is the author's self-citation, Costa (2025), discussed separately in the Conflict of Interest Statement. Nowhere in this paper does an industry report carry the same weight as a tier (i) study.

Inclusion criteria

Empirical studies, field experiments, and industry reports with a stated methodology, as well as peer-reviewed articles.

A focus on generative AI applied to project management processes, consulting, or corporate document automation.

Publication between 2020 and 2026, in English or Portuguese.

Exclusion criteria

Studies of AI in domains with no direct application to PMO or consulting (for example, manufacturing or clinical health).

Promotional material with no stated methodology or data.

Duplicates, and sources whose full text or methods summary was unavailable.

Selection process

Selection followed the PRISMA stages in adapted form: identification of records, screening by title and abstract, full-text reading against the inclusion and exclusion criteria, and final inclusion in the synthesis. Two limitations apply, and this section states them directly rather than leaving them to Section 3.9. First, a single reviewer did the screening, so there is no inter-rater agreement statistic.

Second, no protocol was registered in advance on a platform such as PROSPERO or OSF Registries. Given these limits and the exploratory scope, this paper does not present a quantified PRISMA flow diagram. A count obtained outside licensed indexed databases and without double screening would suggest more precision than the process supports. The full count belongs to the registered systematic review proposed in Section 4.5. The sources analyzed in depth (Section 2 and Appendix A) were chosen for their direct relevance to the research questions, their stated methodological rigor, and the verifiability of the primary source.

Data charting

For each included source, the review charted the following fields, which systematic reviews call data extraction: authors and year; study type; context or population; type of AI intervention; reported metrics; and success factors and limitations stated by the authors. Table 2 in Appendix A reproduces the full chart.

Synthesis and analysis

The thematic synthesis organized findings along five axes: (1) types of automatable deliverables or tasks, (2) reported productivity gains, (3) success factors, (4) risks and limitations, and (5) governance and quality gaps. These axes structure the findings in Section 4.

Study quality assessment

The review applied a simplified quality criterion. It separates studies with a stated experimental design, either randomized or pre-registered, from industry reports and other material without equivalent methodological validation. The latter appear as contextual evidence, not as primary

evidence of causal effects (see Section 3.2.1).

Methodological limitations

Single reviewer: one person selected and assessed the sources, so the review reports no inter-rater agreement (kappa).

No registered protocol (for example, on PROSPERO or OSF Registries), which limits exact replication of the selection process.

Limited search coverage: this phase relied on open-access databases and institutional publications, not on an exhaustive search of Scopus and Web of Science.

Publication bias: successful studies are published and publicized more often than failures or null results.

Heterogeneous metrics: studies report productivity in different units and designs, which rules out statistical meta-analysis and led to thematic synthesis.

Limited external validity: the central experiments rely on specific samples. They include online experiment participants (NOY; ZHANG, 2023; DOSHI; HAUSER, 2024), agents at a single Fortune 500 software company (BRYNJOLFSSON; LI; RAYMOND, 2025), consultants at a single global firm (DELL'ACQUA et al., 2026), and recruiters in a single experiment (DELL'ACQUA, 2022). None of these studies reports results specific to Brazil or other Latin American markets, so effect sizes may not transfer directly to those settings.

A fast-moving field: the formal literature lags practice. One source in this review, Dell'Acqua et al. (2026), moved from working paper to journal article during the review period, and some of its reported figures changed in the published version.

FINDINGS AND THEORETICAL PROPOSITIONS

Productivity gains (experimental evidence) and an illustrative comparison (theoretical proposition)

The experimental literature supports a range of gains, not a single number. The size of the gain depends on the task and on where the task sits relative to the jagged frontier (DELL’ACQUA et al., 2026). In customer support, productivity rose by about 15% (BRYNJOLFSSON; LI; RAYMOND, 2025). In mid-level writing tasks, time fell by about 40% and quality ratings rose by about 18% (NOY; ZHANG, 2023). In consulting tasks inside the frontier, consultants completed 12.2% more tasks, worked 25.1% faster, and produced higher-quality work (DELL’ACQUA et al., 2026). The same body of research supplies counterweights: lower accuracy outside the frontier (DELL’ACQUA et al., 2026), overreliance on automation (DELL’ACQUA, 2022), and less diverse output across users (DOSHI; HAUSER, 2024). Any reported productivity gain should therefore be read as conditional on calibrated human oversight. It is not an automatic property of the technology.

As a comparison from an adjacent domain, Costa (2025) reports a 45% reduction in FP&A cycle time and a 38% gain in forecast accuracy. These figures sit at the top of the range above. The comparison is an illustrative theoretical proposition drawn by analogy, not a validated statistical inference. No study has compared the two domains in one experimental design. The difference in size could reflect a real difference in task structure, or it could reflect uncontrolled differences in context, tool, or sample. No third party has yet replicated Costa (2025). The working hypothesis is that more structured, rule-based deliverables sit more comfortably, and with less risk, inside the frontier than strategic deliverables do. Section 4.5 proposes testing it. See also the Conflict of Interest Statement.

Automation maturity map (the author’s theoretical proposition)

Table 1 is a theoretical proposition. The author deduced it from the task-structure and

recurrence criterion that underlies the jagged frontier concept (DELL'ACQUA et al., 2026). It is not an empirical count taken from the reviewed studies, because none of them measured the five PMO deliverable types listed. Read it as an ordering hypothesis to be tested, not as an observed result.

Table 1 – Hypothesized automation maturity of PMO deliverables

PMO deliverable type	Degree of structure and recurrence	Hypothesized position relative to the frontier	Theoretical basis
Status reports and dashboards	High: fixed format, recurring data	Mostly inside the frontier	Dell'Acqua et al. (2026)
Meeting minutes	Medium-high: fixed format, variable content	Mostly inside the frontier, with light human review	Dell'Acqua et al. (2026)
Schedules and project plans	Medium: rules plus judgment about dependencies	Intermediate zone of the frontier	Dell'Acqua et al. (2026)
Predictive analyses and risk reports	Medium-low: requires interpreting data	Near the edge of the frontier; needs more oversight	Ji et al. (2023); Dell'Acqua et al. (2026)
Strategic documents and first drafts of proposals	Low: specialized judgment and tacit client knowledge	Outside or at the outer edge of the frontier; higher risk of plausible-sounding error	Dell'Acqua et al. (2026); Doshi and Hauser (2024)

Source: Prepared by the author (2026), by deduction from the studies cited in the last column.

The “hypothesized position” column reflects the author’s deduction, not a measurement.

Recurring critical success factors

Four factors recur across the sources reviewed. The first is data quality and standardization. The Deloitte AI Institute (2024b) reports that data problems led 55% of surveyed organizations to avoid certain generative AI use cases. The second is human oversight calibrated to the task. Dell'Acqua et al. (2026) describe “centaur” consultants, who hand discrete tasks to AI, and “cyborg” consultants, who weave AI into their whole workflow. Dell'Acqua (2022) shows what happens when oversight lapses. The third is change management: the Deloitte AI Institute (2024a) identifies talent, governance, and

risk management as the main readiness gaps. The fourth is alignment between automation and the firm's pricing strategy. This paper infers that factor from the Everest Group (2025) account of the shift toward value-led delivery models, not from experimental evidence.

Gap identification

The synthesis confirms that the literature remains fragmented across domains: general professional writing, customer support, broad management consulting, and creativity. It lacks a practical framework, tested in a real multi-client consulting environment, that organizes the automation of PMO deliverables by task type and by the level of human oversight required. This gap points to a second line of research: a governance and quality control framework for AI-generated output in multi-client consulting. The author plans to develop it in a follow-up paper.

Proposed research agenda

Run a full, registered systematic review with a prior protocol (PROSPERO or OSF Registries), a search of Scopus and Web of Science, independent double screening, an inter-rater agreement statistic, and a fully quantified PRISMA flow diagram.

Validate the propositions in Sections 4.1 and 4.2 in a real multi-client consulting environment. Measure cycle time before and after automation, the share of deliverables automated by type, and the rate of human rework. This directly tests whether the size of the gain varies with how structured the deliverable is.

Develop a quality governance and human-in-the-loop framework for generative AI output in multi-client consulting, as a companion study that follows from the risks identified in Section 2.4.

Have third parties replicate the Costa (2025) results in FP&A, so that the comparison in Section 4.1 can stop being merely illustrative.



DISCUSSION AND CONCLUSION

This paper mapped what the literature says about automating PMO deliverables with generative AI in professional services consulting. The margin pressure described in the introduction is not specific to the author's work context. Several independent industry sources document it as a structural trend (EVEREST GROUP, 2025; DELOITTE AI INSTITUTE, 2024a, 2024b; PMI, 2024), although that evidence is descriptive and not causal. Stronger experimental evidence (NOY; ZHANG, 2023; BRYNJOLFSSON; LI; RAYMOND, 2025; DELL'ACQUA et al., 2026) shows that large productivity gains can occur outside the author's own setting. Their size varies with the task.

The paper also gave the evidence against uncritical automation equal space. Overreliance on automation (DELL'ACQUA, 2022) and lower diversity of output (DOSHI; HAUSER, 2024) qualify the productivity story. A firm that scales automation without calibrating human oversight may trade accuracy for speed. It may also deliver strategically similar work to different clients, which undercuts the value of strategy consulting. For PMO leaders and consulting partners, the main lesson is that automation is a calibration decision, not a yes-or-no choice. The right level depends on the type of deliverable and on how structured it is.

One implication concerns the context where this paper originates. This review found no study that reports results specific to Brazil or other Latin American markets. Effect sizes, and the ordering in Table 1, may not carry over to firms that work under different regulatory, contractual, and digital-maturity conditions. That strengthens the case for the local validation proposed in Section 4.5.

Three limitations deserve direct mention. First, this is a scoping review based on open-access sources and institutional publications, screened by a single reviewer without a registered protocol. It is not a full systematic review with an exhaustive search of licensed databases. Second, Costa (2025) is specific to FP&A, no third party has replicated it, and its extrapolation here is a working hypothesis. The Conflict of Interest Statement below addresses this. Third, the metrics differ across studies, which ruled out a statistical meta-analysis and led to a thematic synthesis.

The paper contributes a single set of explicitly labeled theoretical propositions that pull together evidence now spread across separate research areas. It sets evidence for automation against evidence of risk, and it defines a research agenda with two parts: empirical validation in a multi-client setting, and a quality governance framework for generative AI applied to corporate deliverables.

CONFLICT OF INTEREST STATEMENT

This paper cites a case study that the author previously published (COSTA, 2025) on automating FP&A with generative AI, and it uses that study as a comparison point. The author discloses this self-citation so that readers and reviewers can consider possible self-reference bias when they read Sections 1.5 and 4.1. No third party has independently replicated the results in Costa (2025), and applying them to the PMO domain is a working hypothesis, not corroborated evidence.

The author also plans a follow-up paper on a quality governance framework for generative AI in multi-client consulting environments (Section 4.5), which forms part of a continuing research program.

REFERENCES

BRYNJOLFSSON, Erik; LI, Danielle; RAYMOND, Lindsey R. Generative AI at work. *The Quarterly Journal of Economics*, Oxford, v. 140, n. 2, p. 889-942, May 2025. DOI: 10.1093/qje/qjae044.

COSTA, Cesar Lopes Maçol. [Full title to be confirmed by the author]. *Promestre – Journal of Innovation and Management*, [S. l.], 2025. [Volume, issue, pages, and DOI to be confirmed by the author.]

DELL'ACQUA, Fabrizio. *Falling asleep at the wheel: human/AI collaboration in a field experiment on HR recruiters*. Boston: Harvard Business School, Laboratory for Innovation Science, 2022. Working paper.

DELL'ACQUA, Fabrizio et al. Navigating the jagged technological frontier: field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science*, Catonsville, v. 37, n. 2, p. 403-423, Mar./Apr. 2026. DOI: 10.1287/orsc.2025.21838.

DELOITTE AI INSTITUTE. The state of generative AI in the enterprise: now decides next: quarter one report. [S. l.]: Deloitte, Jan. 2024a. Available at: <https://www.deloitte.com/content/dam/assets-shared/docs/services/consulting/2024/gx-state-of-gen-ai-report.pdf>. Accessed on: Oct. 6, 2026.

DELOITTE AI INSTITUTE. The state of generative AI in the enterprise: third quarterly report [press release summary]. [S. l.]: Deloitte, 2024b. Available at: <https://www.deloitte.com/uk/en/about/press-room/deloitte-ai-institute-state-of-generative-ai-in-the-enterprise-report.html>. Accessed on: Oct. 6, 2026.

DOSHI, Anil R.; HAUSER, Oliver P. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, Washington, DC, v. 10, n. 28, art. eadn5290, 12 July 2024. DOI: 10.1126/sciadv.adn5290.

EVEREST GROUP. Market report on the consulting landscape: H1 2025. [S. l.]: Everest Group, 2025. Available at: <https://www.everestgrp.com/report/egr-2025-33-r-7539/>. Accessed on: Oct. 6, 2026.

JI, Ziwei et al. Survey of hallucination in natural language generation. *ACM Computing Surveys*, New York, v. 55, n. 12, art. 248, p. 1-38, Dec. 2023. DOI: 10.1145/3571730.

NOY, Shakked; ZHANG, Whitney. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, Washington, DC, v. 381, n. 6654, p. 187-192, July 2023. DOI: 10.1126/science.adh2586.

PAGE, Matthew J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, London, v. 372, art. n71, 29 Mar. 2021. DOI: 10.1136/bmj.n71.

PROJECT MANAGEMENT INSTITUTE. Project Management Institute leading the transformation of the project management profession with latest next-generation AI offerings [press release]. Newtown Square: PMI, 19 Sept. 2024. Available at: <https://www.pmi.org/about/press-media/2024/pmi-leads-transformation-of-profession>. Accessed on: Oct. 6, 2026.

TRICCO, Andrea C. et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, Philadelphia, v. 169, n. 7, p. 467-473, Oct. 2018. DOI: 10.7326/M18-0850.

