

Generative AI in Management Research: Assistant, Co-Author, or Contaminant?

Ricardo Limongi¹ 

¹ Universidade Federal de Goiás, Faculdade de Contabilidade Economia e Administração, Goiânia, GO, Brazil

How to cite: Limongi, R. (2026). Generative AI in management research: Assistant, co-author, or contaminant?. *BAR-Brazilian Administration Review*, 23(3), e260285.

DOI: <https://doi.org/10.1590/1807-7692bar2026260285>

Keywords:

artificial intelligence; research; ethics; authorship; peer review.

JEL Code:

M10

Publication date:

October 01, 2026.

Corresponding author:

Ricardo Limongi
Universidade Federal de Goiás,
Faculdade de Contabilidade Economia e Administração,
Av. Esperança, s/n, Chácaras de Recreio Samambaia,
CEP 74690-900, Goiânia, GO, Brazil.

Editor-in-Chief:

Ricardo Limongi 
(Universidade Federal de Goiás, Brazil)

Editorial assistants:

Eduarda Anastacio and Simone Rafael (ANPAD, Maringá, Brazil).

ABSTRACT

Generative artificial intelligence has moved from the margins to the center of management research in the space of three years, yet the scholarly community still lacks a shared vocabulary for distinguishing legitimate from illegitimate uses across the research lifecycle. This editorial argues that the technology is neither an assistant to be celebrated uncritically, a co-author to be credited, nor a contaminant to be banned; it is a general-purpose instrument whose legitimacy depends on where it is used, how its outputs are verified, and whether its role is disclosed. We propose a reflexive checklist of five dimensions for deciding whether to use AI at a given stage of research, map seven levels of AI involvement from language polishing to autonomous research agents, and introduce a decision framework organized by the verifiability of AI outputs and their centrality to the scholarly contribution. The editorial concludes with the *Brazilian Administration Review's* explicit expectations for AI-assisted submissions, including a structured disclosure policy, full author accountability, symmetric obligations for reviewers, and a commitment not to use AI-detection software in editorial decisions. Transparency, not prohibition, is the standard the field needs.

Data Availability: BAR – Brazilian Administration Review encourages data sharing but, in compliance with ethical principles, it does not demand the disclosure of any means of identifying research subjects.

Plagiarism Check: BAR maintains the practice of submitting all documents received to the plagiarism check, using specific tools, e.g.: iThenticate.

Peer review: is responsible for acknowledging an article's potential contribution to the frontiers of scholarly knowledge on business or public administration. The authors are the ultimate responsible for the consistency of the theoretical references, the accurate report of empirical data, the personal perspectives, and the use of copyrighted material. This content was evaluated using the double-blind peer review process. The disclosure of the reviewers' information on the first page is made only after concluding the evaluation process, and with the voluntary consent of the respective reviewers.

Copyright: The authors retain the copyright relating to their article and grant the journal BAR – Brazilian Administration Review, the right of first publication, with the work simultaneously licensed under the Creative Commons Attribution 4.0 International license (CC BY 4.0) The authors also retain their moral rights to the article, including the right to be identified as the authors whenever the article is used in any form.



THE PARADOX OF ACCELERATION

In a recent editorial in this journal, we argued that systematic literature reviews had reached an inflection point: an expanding body of reviews that maps more knowledge while generating fewer transformative insights, a tendency that generative artificial intelligence (AI) has not created but has exposed and intensified (Limongi et al., 2026). That diagnosis, we noted, was not really about reviews. It was about a structural condition of contemporary scholarship that reviews merely made visible for the first time. This editorial generalizes the argument: what generative AI did to the systematic review, it is now doing to the entire chain of knowledge production in management research.

The condition is best described as a paradox of acceleration. Producing a manuscript has never been easier. Large language models draft fluent prose, summarize literature in seconds, generate analysis code on demand, and simulate respondents at negligible cost. Controlled evidence confirms substantial productivity gains in professional writing tasks (Noy & Zhang, 2023), and scholars across fields report the same in their own work. Yet the marginal contribution of each additional manuscript has rarely been so uncertain. Grimes et al. (2023) capture the shift precisely: a profession organized around the scarcity of excellent work must now confront its potential abundance, and the assumption that authors, reviewers, and editors will simply act in good faith is inadequate to that transition. What was scarce, polished, publishable output is becoming abundant; what the system took for granted careful verification and accountable judgment is becoming the scarce resource.

The management and business disciplines have not been silent. Editors in information systems have charted the dual, Janus-faced character of the technology for scholarly conduct (Susarla et al., 2023); marketing scholars have mapped its implications for research, teaching, and practice (Peres et al., 2023); multidisciplinary teams have cataloged its opportunities and threats at length (Dwivedi et al., 2023); and management educators have framed it as a set of paradoxes to be managed rather than resolved (Lim et al., 2023). Beyond the field, Messeri and Crockett (2024) issue the deepest warning: AI tools promise productivity and objectivity while exploiting our cognitive limitations, creating illusions of understanding in which we produce more but understand less, and scientific monocultures in which the questions best suited to AI crowd out the rest.

What this growing literature has not yet delivered, and what editors, reviewers, supervisors, and examination committees improvise daily in its absence, is a shared, operational vocabulary for distinguishing legitimate from illegitimate uses of AI across the stages of a research project. The question authors ask is rarely 'is AI good or bad for science?' It is concrete: may I use it to screen this literature, to write this section, to generate these respondents, and if so, under what conditions and with what disclosure? The absence of an answer produces two symmetrical pathologies. Some authors use AI extensively and disclose nothing, fearing desk rejection. Others avoid demonstrably useful tools, fearing stigma. Both behaviors corrode the same asset: trust in the published record. This editorial proposes a vocabulary and states how the *Brazilian Administration Review (BAR)* will apply it.

SHOULD I USE AI AT THIS STAGE? A REFLEXIVE CHECKLIST

In our previous editorial, we offered a reflexive checklist to help decide whether to undertake a systematic review (Limongi et al., 2026). The same discipline of asking before doing applies here, but the unit of analysis changes: the decision to use AI is not made once per project. It is made stage by stage at the literature review, at the design, at the analysis, at the writing, and a use that is legitimate at one stage may be illegitimate at the next. Before delegating any research task to a generative system, we suggest researchers interrogate five dimensions.

(1) Verifiability: Can I check what the machine gives me?

- Independent verification: Can I verify every output the AI produced every reference, every number, every claim against a source that does not depend on the AI itself?
- Error detection: If the system hallucinated here, would I catch it? Fluency is not accuracy, and the most dangerous errors are the ones embedded in confident prose.

(2) Intellectual ownership: Is the contribution still mine?

- Origin of the argument: Did the core problematization, theorizing, and interpretation originate with me, or am I curating machine output?

- The viva test: Could I defend every sentence of this manuscript, without notes, before a doctoral committee?

(3) Transparency: Am I prepared to say so?

- Disclosure readiness: Am I prepared to state explicitly what I used, where I used it, and how I verified it?
- The embarrassment heuristic: If disclosure would embarrass me, the use was probably illegitimate.

(4) Epistemic risk: What does this use do to the knowledge itself?

- Bias and contamination: Does this use import the biases of training data into my evidence, or create circularity AI summarizing AI-written literature, AI reviewing AI-drafted text?
- Monoculture pressure: Am I choosing this question or method because it matters, or because it is the one AI handles well (Messerli & Crockett, 2024)?

(5) Competence: Do I know enough to supervise?

- Evaluative capacity: Do I understand the task well enough to evaluate the output, or am I delegating judgment I do not possess?
- The apprentice rule: A tool that replaces a skill I never acquired is not assistance; it is substitution without supervision.

These five dimensions convert an anxious, binary question to use or not to use into a structured professional judgment. They also make explicit what has always been true of research instruments, from statistical software to survey panels: legitimacy is a property of the use, not of the tool.

A PORTFOLIO OF AI USES ACROSS THE RESEARCH LIFECYCLE

Just as the systematic review is not one method but a portfolio of approaches with different theoretical ambitions (Limongi et al., 2026), ‘using AI in research’ is not one practice. It is a spectrum of involvements that differ in the depth of delegation, the conditions under which they are legitimate, and the disclosure they demand. Table 1 organizes seven levels of that spectrum, from the most superficial to the most consequential.

Table 1. A portfolio of generative AI uses in management research.

Use	Central question	Legitimacy conditions	Disclosure expectation	Key risks
Language polishing and translation	Is my prose clearer and more accessible?	Universal, provided meaning is preserved and checked by the authors	General statement	Homogenization of scholarly voice; loss of authorial style
Literature discovery and screening	What relevant work am I missing?	Every retrieved source independently verified against the original	General statement in methods	Hallucinated or misattributed references; coverage bias
Coding and analysis assistance	Is my analytical pipeline correct and efficient?	Human validation of logic, code, and outputs against known benchmarks	Methods section	Silent computational errors; unexamined analytical defaults
Ideation and structured critique	Where is my argument weak or underdeveloped?	Human retains full interpretive and theoretical judgment	Optional	Sycophantic feedback; convergence toward average ideas
Drafting of text sections	Who is the intellectual author of this passage?	Substantial human revision; argument and claims originate with the authors	Explicit, section-level disclosure	Authorship dilution; fabricated claims embedded in fluent prose
Synthetic data and silicon samples	Can model-generated respondents stand in for humans?	Validation against human benchmarks; explicit reporting of prompts, models, and parameters	Full protocol disclosure	Construct invalidity; training-data bias presented as evidence
Autonomous research agents (end-to-end pipelines)	Is this still my research?	Rarely legitimate at present; every stage must remain auditable and human-approved	Full protocol disclosure	Loss of accountability; error propagation across stages

Note. Elaborated by the authors. Legitimacy conditions and disclosure expectations synthesize guidance in [Committee on Publication Ethics. \(2023\). Authorship and AI tools: COPE position statement.](#) <https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools>, [International Committee of Medical Journal Editors. \(2025\). Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals.](#) <https://www.icmje.org/recommendations/>, [Sarstedt, M., Adler, S. J., Rau, L., & Schmitt, B. \(2024\). Using large language models to generate silicon samples in consumer and marketing research: Challenges, opportunities, and guidelines. *Psychology & Marketing*, 41\(6\), 1254–1270.](#) <https://doi.org/10.1002/mar.21982>, and [Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. \(2023\). The Janus effect of generative AI: Charting the path for responsible conduct of scholarly activities in information systems. *Information Systems Research*, 34\(2\), 399–408.](#) <https://doi.org/10.1287/isre.2023.ed.v34.n2>.

Language, discovery, and analysis: Assistance under verification

The first three levels are where most researchers already operate. Language polishing is the least controversial use and, for a journal like *BAR* published in English from a country where English is not the first language, arguably the most democratizing: it narrows a gap that has long taxed scholars from the Global South, who pay in time, money, and rejection rates for prose that native speakers produce for free. Literature discovery and analysis assistance are equally legitimate but categorically riskier, because their failure modes are silent. A hallucinated reference looks like a real one; a subtly wrong regression script runs without complaint. The governing principle across these levels is verification: the researcher, not the model, warrants every source and every number that reaches the manuscript (van Dis et al., 2023).

Ideation, critique, and drafting: The authorship boundary

The middle levels press against the meaning of authorship. Using a model as a sparring partner to stress-test an argument, surface counterexamples, or simulate a hostile reviewer leaves intellectual ownership intact and requires little disclosure, though researchers should be alert to the tendency of these systems to produce agreeable, average feedback. Drafting is different. Here the international consensus is unambiguous: AI tools cannot be authors, because authorship requires the capacity to take responsibility for the work, and they have none (Committee on Publication Ethics [COPE], 2023; International Committee of Medical Journal Editors [ICMJE], 2025; Thorp, 2023). Text generation is therefore legitimate only when the argument originates with the authors, the drafted material is substantially revised, and the use is explicitly disclosed. The test is not whether a machine touched the words but whether a human can be held accountable for every claim they carry.

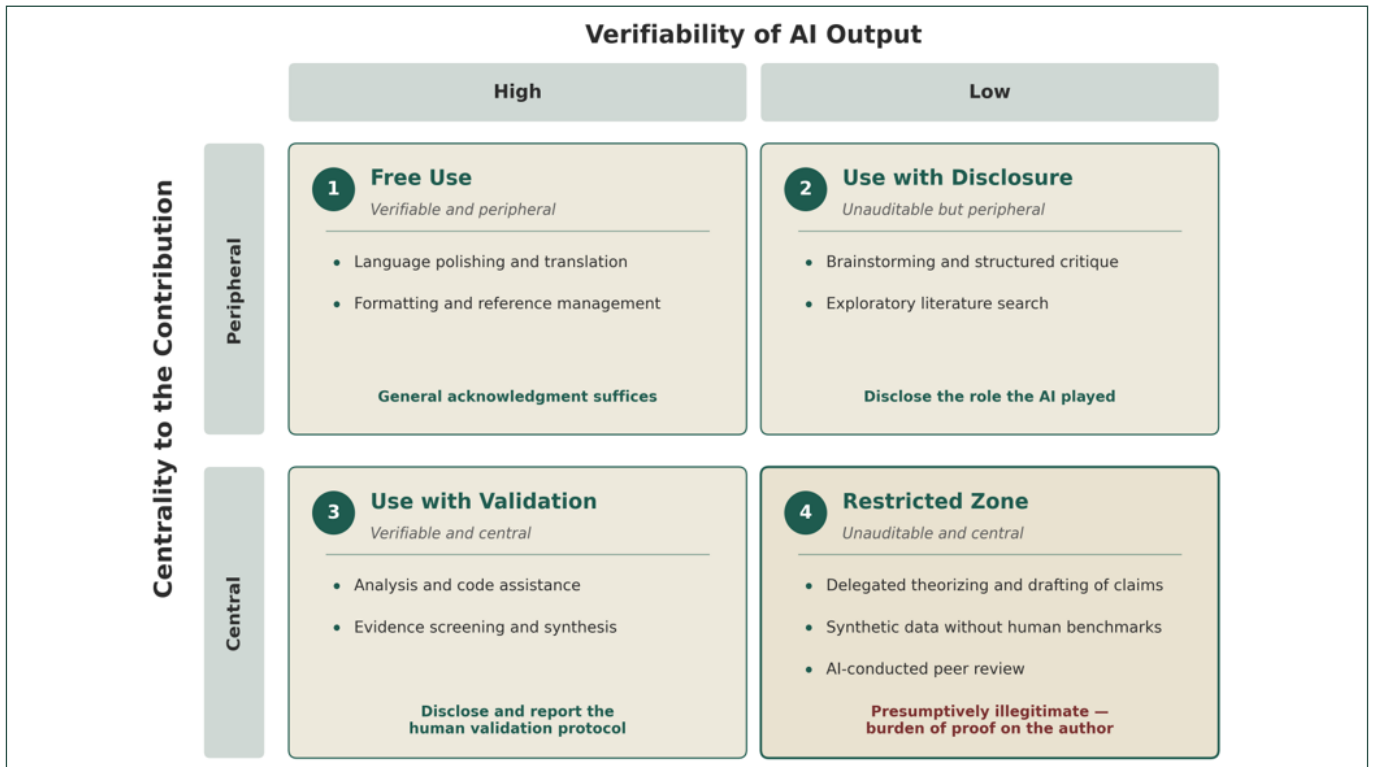
Synthetic data and autonomous agents: The research frontier

The final levels are the frontier on which the next decade of methodological debate will be fought. Large language models can simulate human respondents, so-called silicon samples, with sometimes striking fidelity to human response patterns (Argyle et al., 2023), and a fast-growing literature in consumer and marketing research is mapping where that fidelity holds and where it collapses (Sarstedt et al., 2024). The emerging consensus is conditional: synthetic respondents hold promise upstream, in pretesting and pilot work, but results vary considerably across domains, and their use as primary evidence demands validation against human benchmarks and full disclosure of models, prompts, and parameters (Bail, 2024; Sarstedt et al., 2024). Autonomous agents that chain the entire pipeline search, synthesis, analysis, writing push the question to its limit: at some point, delegation stops being assistance and becomes ghost science, research for which no human can honestly claim responsibility. *BAR* does not consider that point to be safely defined and treats end-to-end automation as presumptively illegitimate.

A DECISION FRAMEWORK

The portfolio describes what researchers do; it does not yet tell them how to decide. Two questions, taken together, do most of that work. The first is verifiability: can the AI's output be independently checked against ground truth, a source, a benchmark, a rerun of the code? The second is centrality: how close does the output sit to the scholarly contribution itself to the theorizing, the evidence, the claims on which the paper stands or falls? Crossing the two yields the framework in Figure 1.

Quadrant 1 (high verifiability, peripheral) is the zone of free use: polishing a paragraph or formatting references neither threatens the contribution nor escapes checking. Quadrant 2 (low verifiability, peripheral) is the zone of disclosure: brainstorming and exploratory search cannot be fully audited, but because they feed rather than constitute the contribution, honest acknowledgment suffices. Quadrant 3 (high verifiability, central) is the zone of validation: analysis code and evidence synthesis sit at the heart of the paper, so legitimacy rests on a documented protocol of human checking, not on trust. Quadrant 4 (low verifiability, central) is the restricted zone: outputs that cannot be audited and that carry the contribution itself delegated theorizing, synthetic evidence without human benchmarks, machine-conducted review. Here the burden of proof lies entirely with the author, and the default answer is no. Like the framework in our previous editorial, this one is indicative rather than mechanical: its purpose is to make the judgment explicit, justified, and inspectable, replacing silence with reasons.



Source: Elaborated by the authors. Quadrant labels indicate the default governance regime; boundaries are indicative, and movement toward the restricted zone always increases the burden of validation and disclosure.

Figure 1. Decision framework for the use of generative AI in management research.

EXPECTATIONS FOR THE BAR

As editors of the *Brazilian Administration Review*, we use this editorial, as we do for systematic reviews, to make the journal's expectations explicit. *BAR* will evaluate AI-assisted submissions against the following five commitments:

Structured disclosure. Manuscripts that used generative AI beyond basic language polishing must include an AI Use Statement in the methods section or a dedicated section specifying which tools and versions were used, at which stages, for which tasks, and how outputs were verified. A generic sentence 'ChatGPT was used to improve the text' is not a disclosure; it is a label. Disclosure will never, by itself, count against a submission.

Accountability. Following [COPE \(2023\)](#) and [ICMJE \(2025\)](#), AI tools cannot be listed as authors, and the human authors are fully responsible for every part of the manuscript, including the parts a machine helped produce. 'The model wrote it' will never be an accepted account of error, fabrication, or plagiarism.

Verification of central uses. Uses falling in the validation and restricted quadrants of Figure 1 analysis, synthesis, and synthetic data must be accompanied by a description of the human validation protocol: what was checked, against what benchmark, by whom. For silicon samples specifically, *BAR* expects validation against human data and full reporting of models, prompts, and parameters ([Sarstedt et al., 2024](#)).

Reviewer symmetry. Obligations do not stop with authors. Reviewers may not upload manuscripts, in whole or in part, to public AI systems: doing so breaches the confidentiality on which peer review rests, and delegated judgment is no longer peer review ([Hosseini & Horbach, 2023](#)). Reviewers who use AI in permitted, limited ways as language support in drafting their own reports must disclose it to the editor.

No AI-detection policing. *BAR* will not use AI detection software as a criterion for editorial decisions. The evidence is clear that such detectors are unreliable and systematically biased against non-native English writers ([Liang et al., 2023](#)), which is to say, against precisely the scholarly communities this journal exists to serve. Our governance rests on disclosure, accountability, and verification, not on unreliable forensics.

These commitments are deliberately symmetric: they bind authors, reviewers, and editors alike, and this editorial binds us first. In that spirit, and consistent with the policy it announces, the authors disclose their use of AI tools in preparing this text in the statement below.

ASSISTANT, CO-AUTHOR, OR CONTAMINANT?

Returning to the question in our title. Is generative AI an assistant? When its outputs are verified and its role disclosed, yes, and for scholars writing science in a second language, a consequential one. Is it a co-author? Never: authorship is accountability, and accountability cannot be delegated to a system that bears none. Is it a contaminant? Only under one condition: when we pretend we are not using it. Undisclosed use converts a legitimate instrument into a corrosive one, not because the text is worse, but because the reader can no longer tell what has been checked, by whom, and who answers for it.

Our previous editorial closed by observing that, for systematic reviews, the method was not the problem; the ambition was (Limongi et al., 2026). The parallel conclusion imposes itself here. The technology is not the problem. The silence about it is.

REFERENCES

- Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3), 337–351. <https://doi.org/10.1017/pan.2023.2>
- Bail, C. A. (2024). Can Generative AI improve social science? *Proceedings of the National Academy of Sciences*, 121(21), e2314021121. <https://doi.org/10.1073/pnas.2314021121>
- Committee on Publication Ethics. (2023). Authorship and AI tools: COPE position statement. <https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Grimes, M., von Krogh, G., Feuerriegel, S., Rink, F., & Gruber, M. (2023). From scarcity to abundance: Scholars and scholarship in an age of generative artificial intelligence. *Academy of Management Journal*, 66(6), 1617–1624. <https://doi.org/10.5465/amj.2023.4006>
- Hosseini, M., & Horbach, S. P. J. M. (2023). Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Research Integrity and Peer Review*, 8, Article 4. <https://doi.org/10.1186/s41073-023-00133-5>
- International Committee of Medical Journal Editors. (2025). Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. <https://www.icmje.org/recommendations/>
- Liang, W., Yuksekogonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), 100779. <https://doi.org/10.1016/j.patter.2023.100779>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790. <https://doi.org/10.1016/j.ijme.2023.100790>
- Limongi, R., Bassi-Suter, M., & Galleli, B. (2026). Systematic literature reviews in business administration: Dead, alive, or in need of reinvention? *Brazilian Administration Review*, 23(2), e260146. <https://doi.org/10.1590/1807-7692bar2026260146>
- Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002), 49–58. <https://doi.org/10.1038/s41586-024-07146-0>
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192. <https://doi.org/10.1126/science.adh2586>
- Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*, 40(2), 269–275. <https://doi.org/10.1016/j.ijresmar.2023.03.001>
- Sarstedt, M., Adler, S. J., Rau, L., & Schmitt, B. (2024). Using large language models to generate silicon samples in consumer and marketing research: Challenges, opportunities, and guidelines. *Psychology & Marketing*, 41(6), 1254–1270. <https://doi.org/10.1002/mar.21982>
- Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. (2023). The Janus effect of generative AI: Charting the path for responsible conduct of scholarly activities in information systems. *Information Systems Research*, 34(2), 399–408. <https://doi.org/10.1287/isre.2023.edv34.n2>
- Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313. <https://doi.org/10.1126/science.adg7879>
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7>

Author

Ricardo Limongi 

Universidade Federal de Goiás, Faculdade de Contabilidade Economia e Administração
Rua Samambaia, s/n, Chácaras Calífornia, CEP 74001-970, Goiânia, GO, Brazil
ricardolimongi@ufg.br