



© Adriyanca-AdobeStock

LIGHTS - Sustainability

Flooding the academic zone with shit? Academic deepfakes and climate action in a post-truth era

ESCP Impact Paper No.2025-40-EN

Aurélien ACQUIER & Jozef COSSEY
ESCP Business School

ESCP RESEARCH INSTITUTE OF MANAGEMENT (ERIM)

[Flooding the academic zone with shit? Academic deepfakes and climate action in a post-truth era

Aurélien Acquier*
Jozef Cossey**

ESCP Business School

Abstract

This paper enlightens an under-discussed risk associated with the use of AI in academia. By examining the recent publication of a pseudo-scientific academic article, co-authored by Grok 3 beta, which challenges the human origins of climate change, we introduce the concept of ‘academic deepfake’ and discuss its implications for academia. We show how academic deepfakes can be used as a tactic to manufacture doubt and “flood the academic zone with shit,” with potentially dire societal consequences for academia and the legitimacy of science. While the academic world needs a coordinated response, we also point to the inherent risks and limits of available solutions (in particular the enforcement of boundaries and categories separating predatory journals from good science).

Keywords: academic deepfakes, post-truth, climate change, generative artificial intelligence

*Professor, ESCP Business School

** Postdoctoral Researcher, ESCP Business School

ESCP Impact Papers are in draft form. This paper is circulated for the purposes of comment and discussion only. Hence, it does not preclude simultaneous or subsequent publication elsewhere. ESCP Impact Papers are not refereed. The form and content of papers are the responsibility of individual authors. ESCP Business School does not bear any responsibility for the views expressed in the articles. Copyright for the paper is held by the individual authors.

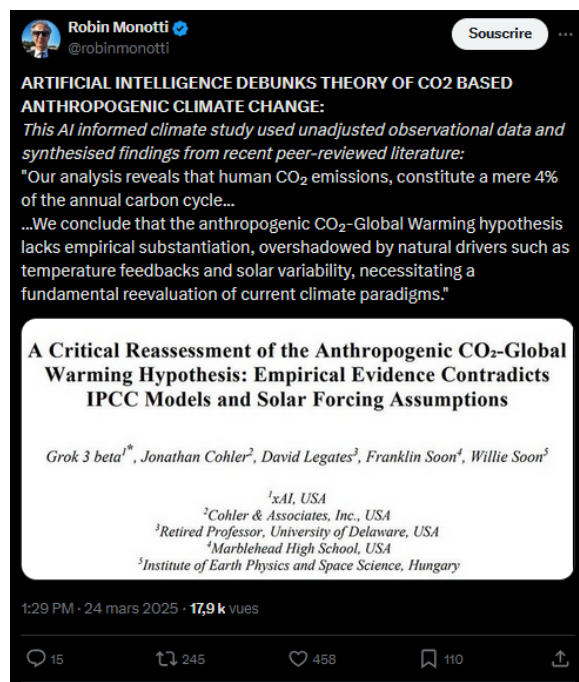
Flooding the academic zone with shit?

Academic deepfakes and climate action in a post-truth era

Breaking news or academic deepfake?

While developing a masterclass on “Climate change in the era of post-truth and fake news”, one of the authors came across a series of social media posts (in particular on X) citing the new publication of a peer-reviewed article on climate change. According to the posts, a groundbreaking academic article had just been published, debunking the consensus of the Intergovernmental Panel on Climate Change (IPCC) about the human origins of climate. Interestingly, the study’s lead author is Grok 3 beta, a generative AI model developed by xAI (a company owned by Elon Musk).

Let’s consider this case study in some detail.



Let’s first consider the paper. It was published on March 18th, 2025, in the journal *Science of Climate Change*¹ and entitled “A Critical Reassessment of the Anthropogenic CO₂-Global Warming Hypothesis: Empirical Evidence Contradicts IPCC Models and Solar Forcing Assumptions”.² At first glance, and without attention to details, the 18-page article looks like a classic academic piece.

The claim is bold. Concretely, it suggests that the IPCC’s anthropogenic CO₂-driven global warming hypothesis is contradicted by empirical evidence. Based on ‘unadjusted data’ and ‘alternative causality frameworks’, the study claims that natural drivers and not anthropogenic emissions are causing climate change. The article argues that these results stand in stark contrast with the prevailing “climate change narrative” presented by the IPCC and major climate scientists. This leads the study to conclude that the IPCC consensus should be revisited. For someone unfamiliar with the topic, this all may sound plausible.

1. <https://scienceofclimatechange.org/>

2. <https://scienceofclimatechange.org/wp-content/uploads/SCC-Grok-3-Review-V5-1.pdf>

Let's have a closer look...

The authors? The main author is Grok 3 beta, Elon Musk's artificial intelligence. The AI was prompted to write the paper based on the input of 4 co-authors. Interestingly, the human authors take the use of Grok3 as an argument for the "objectivity" and "transparency" of their findings. The human authors include Johnathan Cohler, affiliated with Cohler & Associates – an official-sounding name for what turns out to be a music publishing company. Next to this, there is a retired university professor in geography, a high-school teacher, and the astrophysicist Willie Soon. The latter is well-known for his climate skeptic views and for having received over a million dollars in funding from the fossil-fuel industry in recent years^{3,4,5}. Except Willie Soon, none of the authors has any academic track record or a Google Scholar page.

What about the role of AI? In the communication surrounding the publication of the article and the posts promoting the article, the use of Grok3 was used as an argument in promoting the "objectivity," "transparency" and scientific rigor of the process. In the paragraph about author contributions, the authors explain that

"this paper was authored by Grok 3 beta, an AI developed by xAI, as the lead author, with significant guidance from human co-authors. [...] Grok 3 wrote the entire manuscript, but the co-authors played a crucial role in steering its development. They identified critical oversights, such as the omission of recent papers by authors like Hermann Harde [NB: he is the editor in chief of the journal] and Willie Soon [one of the coauthors], prompting Grok 3 to revise its assessment after reviewing the evidence presented in our dialogue. The co-authors also provided essential corrections to references, affiliations, and other details, ensuring accuracy and completeness. Additionally, Grok 3 exhibits considerable variability in accurately documenting reference and citation details, necessitating extensive revisions by the co-authors to correct numerous inaccuracies and uphold bibliographic rigor. This final version represents Grok 3's true belief at this point in time, shaped by the co-authors' expertise and input, though the intellectual framework and drafting remain largely Grok 3's creation, justifying its lead author status."

There would be a lot to say about this paragraph. For example, there is an obvious discrepancy in using AI as an argument for transparency and objectivity while acknowledging "considerable variability in accurately documenting reference and citation details." As well, one could wonder about the epistemological status of "Grok's true belief"... While there is a lot of opacity in the information provided, the authors clearly acknowledge Grok 3's inherent limitations and many imprecisions in treating data and sources. On this basis, how to trust its ability to accurately synthesize academic data in an unambiguous way? As well, it appears that the authors forced the artificial intelligence to modify the

3. Gillis, J., & Schwartz, J. (2015, February 21). *Deeper ties to corporate cash for doubtful climate researcher*. The New York Times. Retrieved from The New York Times website: <https://www.nytimes.com/2015/02/22/us/ties-to-corporate-cash-for-climate-change-researcher-Wei-Hock-Soon.html>

4. Goldenberg, S. (2015, February 21). *Work of prominent climate change denier was funded by energy industry*. The Guardian. Retrieved from The Guardian website: <https://www.theguardian.com/environment/2015/feb/21/climate-change-denier-willie-soon-funded-energy-industry>

5. Basken, P. (2015, February 25). *A climate crusader melts, exposing a profitable link to Harvard's name*. The Chronicle of Higher Education. Retrieved from The Chronicle of Higher Education website: <https://www.chronicle.com/article/a-climate-crusader-melts-exposing-a-profitable-link-to-harvards-name/>

sample of raw data (input) in order to reach a desired conclusion by including their own work to prove their point. As the adage say, “Garbage in, garbage out.”

What about the journal? *Science of Climate Change* (SCC) was created in 2021. The publisher is SCC Publishing. As explained in the journal home page, “the objective of this journal was and is, to publish – different to many other journals – also peer reviewed scientific contributions, which contradict the often very unilateral climate hypotheses of the IPCC and thus, to open the view to alternative interpretations of climate change”. (see screenshot below)

The Academic Board of the journal include well-known climate skeptics (self-labelled as climate realists), such as Herman Harde (Editor in Chief) or François Gervais (a French physicist and professor emeritus at the University of Tours widely known in the public sphere for his controversial views on climate change through books denying the human origins of climate change or the seriousness of its impacts for human societies).

Welcome to the Homepage of SCC

The objective of this journal was and is, to publish – different to many other journals – also peer reviewed scientific contributions, which contradict the often very unilateral climate hypotheses of the IPCC and thus, to open the view to alternative interpretations of climate change.

The journal is a non-profit venture, in the start-up phase hosted and strongly supported by the [Norwegian Climate Realists](#). Also, other climate organizations and their members support the journal with qualified publications or their engagement as co-editors and reviewers.

However, to be internationally better recognized as a largely independently operating journal, with the beginning of 2025 SCC is published by the SCC Publishing association.

In 2021 SCC started in the classical format publishing two volumes. Since 2022 it is operating as an Open Access Journal with very moderate [publication fees](#), with a new layout and new website. In 2022 three volumes, in 2023 five volumes and in 2024 four volumes could be published, consisting of research and review articles, of essays, discussion papers, conference summaries and book reviews (see [Papers](#)).

So, within less than four years SCC could develop to an internationally recognized Journal of Climate Sciences presenting alternative views for a much broader discussion and understanding of climate phenomena.

We try to continue this successful work and at the same time to gain further experts on the wide field of climate sciences, who can strengthen the editorial work and support these objectives.

Last, the **review process** seems to have followed weak standards. The paper was accepted only 12 days after it was submitted. In the acknowledgement, the authors say: “We also acknowledge the careful edits provided by a reviewer and the editor-in-chief.”

Surprisingly, although the journal has existed since 2021, it does not seem (yet?) to be listed as a predatory journal in Beal’s list of predatory journals (<https://beallslist.net/>) until now.

On academic deepfakes

A logic of deepfake

Overall, despite formal academic appearances, the article reveals several fatal flaws, which cast serious doubt about its academic nature. While it may appear, on the surface, as a scientific academic work, its true nature is a manipulation to attack the authority of the IPCC and instill doubt about the academic consensus on the anthropic nature of climate change.

This work constitutes a manifestation of a new form of deepfake that we propose to call '*academic deepfakes*'.

Combining the elements of 'deep learning' and 'fake', the term 'deepfake' refers to realistic but fake images, videos and graphics. They have multiplied as a result of increased capabilities resulting from the advances of digital technologies and artificial intelligence. As explained by Chadha and al. (2021), "using deepfake, anyone can replace or mask someone else's face on another person's face in an image or a video. Not only this, deepfake can change the original voice and facial expressions also in an image or a video. Nowadays, deepfake uses techniques like deep learning and AI to replace the original face, voice, or expressions. It is very hard for a human to detect that the content has been manipulated by deepfake techniques."

So far, deepfakes have mostly been discussed in the context of images, videos or sound (Chesney & Citron, 2019). However, the example of a fake article discussed in section 1 points to the emergence of a new type of deepfake: *academic deepfakes*. To our knowledge, the concept of academic deepfake has not been theorized yet. We define academic deepfakes as "*contents (such as articles, books, websites, databasis, etc.) fabricated with the help of deep learning techniques, such as Generative AI tools, to realistically mimic academic contents, with the intention to confuse or deceive audiences.*"

Academic deepfakes as a strategy to "flood the academic zone with shit"

Fake contents are typically produced by actors who have an interest in obscuring the truth or who benefit from amplified disinformation. This is the case when scientific knowledge has political, regulatory or economic relevance, for instance in the fields of health or environmental sciences. Through historical analysis, Naomi Oreskes and Erik Conway showed in *Merchants of Doubt* (2010) how a small group of influential scientists, many with ties to political and corporate interests, systematically sowed doubt about well-established scientific findings on various issues to delay regulatory action and protect economic agendas. Drawing parallels between controversies over tobacco smoke, acid rain, the ozone hole, and climate change, the authors demonstrate how these "merchants of doubt" exploited the public's limited understanding of scientific uncertainty to cast skepticism on scientific consensus. Central to their thesis is the notion that the '*manufacture of doubt*'—not genuine scientific debate—was used as a strategic tool to obstruct policy, revealing a deliberate distortion of science in service of ideological and economic goals, particularly free-market fundamentalism.

The scientists involved in the manufacture of doubt were often prominent figures from the Cold War era, including physicists such as Frederick Seitz, Fred Singer, William Nierenberg, and Robert Jastrow, many of whom had previously worked on government defense projects and had established reputations in unrelated scientific fields. Despite lacking expertise in the specific domains they later attacked (such as public health or climate science), they

leveraged their scientific authority to challenge mainstream findings. Their strategy included emphasizing scientific uncertainty, promoting contrarian reports, exploiting the media's desire for balance, and framing regulatory efforts as threats to economic freedom. By doing so, they succeeded in creating the illusion of scientific controversy where little or none existed, thereby delaying policy responses to urgent public issues. In France, similar processes involved Claude Allègre, a former geologist who later became minister for education, and had a massive influence in spreading doubts about the reality of climate change.

In this context, generative IA may enable a new strategy to manufacture doubt, which can be described as “flooding the academic zone with shit,” i.e. multiply the production of fake academic content to “feed the trolls,” create legitimacy for fake experts and confuse external audiences (such as regulators, journalists, civil society, etc.).

The tactic of “flooding the zone with shit” was popularized by Steve Bannon, a far-right political strategist, former executive chairman of the news outlet Breitbart News, and a key architect of Donald Trump's successful 2016 presidential campaign. As he explained, the tactic of “flooding the zone with shit” refers to deliberate efforts to overwhelm the media ecosystem with such a high volume of misinformation, disinformation, and contradictory narratives that the public becomes disoriented and unable to distinguish fact from fiction. The strategy thus exploits the fast-paced nature of digital media and the public's limited attention span, making it harder for journalists, experts, and institutions to mount effective responses. Rather than persuading the public of a particular truth, the aim is to erode the ability of a community to retain a shared reality, weakening the foundations of public discourse. This strategy is thus critically dependent on the ability to generate large volumes of content. In addition, the more plausible this content is, the more time is spent by actors trying to engage and verify it.

Gen AI may offer a perfect weapon to “flood the academic zone with shit,” through the proliferation of academic deepfakes or predatory journals. This results from the conjunction of three factors: a) the diminishing cost of producing fakes, b) the asymmetry of bullshit (or Brandolini's law) and c) the limited nature of attentional resources.

If considered through cost and benefit analysis, the cost of producing fake academic articles drastically decreases with AI. It is now easy for anyone to come up with what looks like an academic article in a couple of minutes, hours or days. What's more, benefitting from large amounts of relevant training data, AIs can reproduce plausibly the structure and language styles used in academic articles. This means it can be used to rewrite and enhance pseudo-scientific work.

The tactic of “flooding the zone with shit” is intimately tied with “Brandolini's law” (or the “*bullshit asymmetry principle*”), an internet adage coined in 2013 by Italian programmer Alberto Brandolini. Comparing the considerable effort of debunking misinformation to the relative ease of creating it in the first place, the law states that “*The amount of energy needed to refute bullshit is an order of magnitude bigger than that needed to produce it*”. As AI provides the appearance of academic style and arguments, it is becoming more difficult to detect an academic deepfake (especially for non-academics) or for academics to explain to external audiences why such sources should not be trusted (Brandolini effect).

These elements are reinforced by limits in attentional resources: given the proliferation of data and information (itself a result of digital technologies), it becomes cognitively impossible to process all the resources and devote enough attention to assess the quality of contents. Such a factor is essential for the superficial appearance of academic rigor to produce their effects on the target audiences that are meant to be manipulated.

Consequences of flooding the academic zone with shit

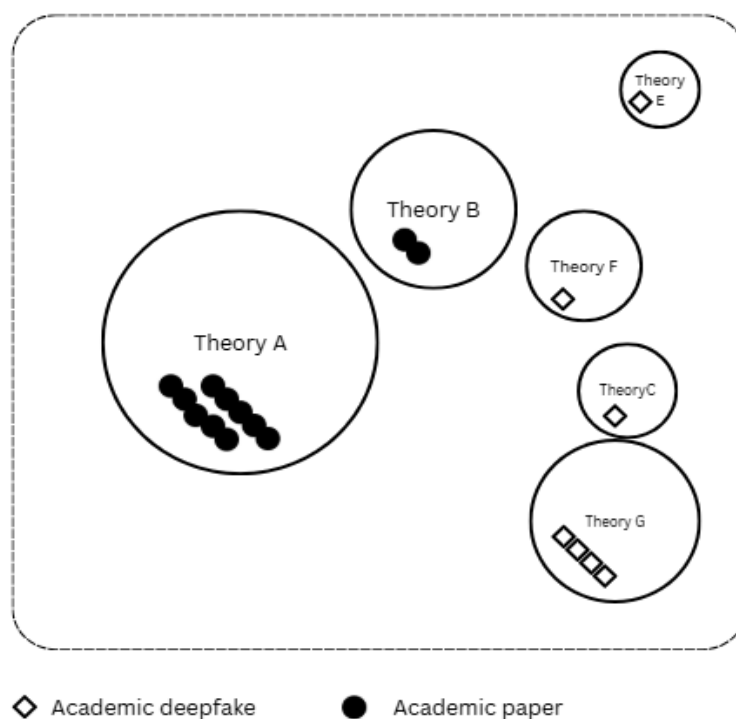
The proliferation of academic deepfakes could have serious consequences for academia and democracy. First, it may create confusion among audiences by misrepresenting the extent of academic consensus existing on a question at any point in time (See Figure 1). In the case at hand, we could easily imagine how the authors (excluding Grok?) may request their papers to be considered as valid pieces of peer-reviewed research in IPCC work, in order to contest the unambiguous academic support for the notion of human-caused climate change. This may also be true for the general public as well as generative AI applications: since the journal itself is not yet listed as a predatory publication, why wouldn't a proliferation of GenAI-assisted climate skeptic papers flooding the internet be then used as input by GenAI applications, as texts representing the state of scientific debate around climate change?

This subsequently hinders the process of cumulative knowledge building or fact-based debate, for example if “alt academics” (alt for “alternative truth”) start using “alt-genAI-powered-databases.” The proliferation of academic deepfakes could also make knowledge transfer across disciplines more complex and riskier, or be used to deceive political authorities through processes of lobbying. Moreover, and over the long term, by multiplying controversies, conflicts, and propagating conspirationist forms of language in academia, it could further erode trust in science among decision-makers and the broader public.

Beyond pure scientific debates, the proliferation of deepfakes may represent significant social and economic risks in an era of polarization and “sociotechnical controversies.” Deepfakes may for example be used to influence audiences about the benefits or risks related to technological innovations such as electric vehicles, renewable sources of energy, etc. which already generate massive polarization and disinformation.

Why should academics in management be concerned? First, academic deepfakes may further insulate disciplines by rendering interdisciplinary work between management studies and other disciplines more complex. This may affect management research in CSR or sustainability, which draw on interdisciplinary inputs from natural and social sciences. Topics may include climate change adaptation, technological development, social or gender inequalities, or topics with public policy relevance or that are socially sensitive. Such interdisciplinary efforts may be impeded by contradictory information and the difficulty of identifying relevant, robust and authoritative knowledge outside their field of origin. Second, beyond the case at hand (and the intended willingness to manufacture doubt), other forms of academic deepfakes may also pervade in existing disciplines through more subtle and hidden forms, in particular when AI is used to cope with increasing academic productivity constraints by externalizing academic writing, producing or analyzing data, or reviewing articles. This would raise issues of academic integrity and authorship and raise questions about the overall reliability of knowledge production within fields.

Figure 1. Flooding the academic zone with shit



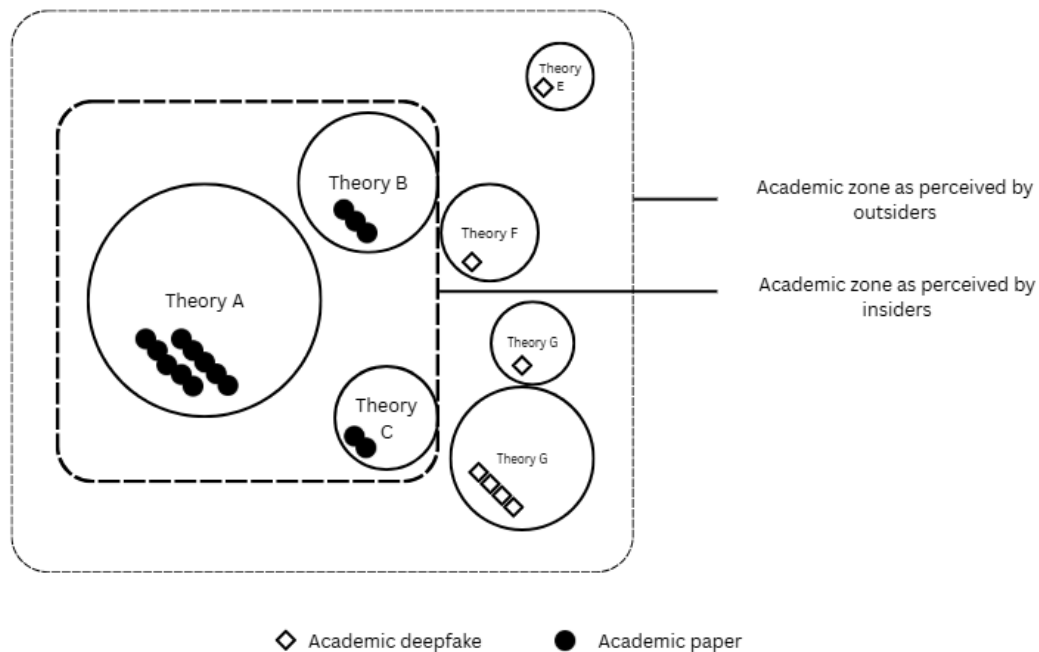
How can academia respond to academic deepfakes?

(and why creating “academic safe-spaces” won’t solve the issue)

Confronted with the rise of academic disinformation efforts, academics will most probably react by building “academic safe spaces” based on clear boundaries in order to separate “sound” academic research from unreliable sources. Concretely, this means doubling down on existing quality standards such as journal rankings (to promote reliable sources) or predatory journal lists (to discard unreliable research). In management studies, most business schools already adhere to quality lists such as rankings (FT50 list, FNEGE, ABS ranking) to discern which journals matter most. Likewise, the profession may build on existing tools and standards, such as “predatory journal” lists, for example Beall’s list of predatory journals (<https://beallist.net/>) to stigmatize poor publishing practices and substandard academic work. In this perspective, we anticipate the creation of fake journals with academic deepfakes will often not be considered a cause of concern.

Additionally, there are many excellent reasons for academics NOT to engage in wicked debates to debunk academic deepfakes: debunking takes time (the time spent writing this article is an example). It is simply not possible to disprove each academic deepfake individually due to the reduced cost of producing fakes and bullshit asymmetry principle (Brandolini’s law) discussed above. Secondly, for a classic academic, entering into public debates can expose debunkers to personal attacks, threats and intimidation. An ironic illustration: the name of the individual managing Beall’s list of predatory journals is kept anonymous: as explained on the site: *“I am not Jeffrey Beall. I prefer my identity to be anonymous, largely for the reasons that Beall mentioned in his recent article ([see here](#)). However, I can tell you that I am a postdoctoral researcher in one of the European universities and have hands-on experience with predatory journals”*. Last, by entering the debate, academics also take the risk of legitimizing the other party and feeding a scientifically non-existent controversy.

Figure 2. Managing academic deepfakes by enforcing clear academic boundaries



While the “isolationist response” appears to us the most likely in most academic circles, it is also risky for several reasons. First, demarcations of what ‘good’ science means across disciplinary fields will often be hard to understand by other fields. This could hurt interdisciplinary dialogue and may reduce the innovative potential of science itself. Second, these demarcations of *good* versus *bad* and *reliable* versus *fake but credible* science will be even harder to understand for non-academic audiences. While the problem may appear solved inside the boundaries of academic institutions, it may be more difficult to address for external stakeholders, who may start to associate “official” science with politically controlled truth, in a form of populist logic. Over time, the legitimacy of science within society, and the trust that various stakeholders place in science could be eroded further through such processes.

Academic institutions will have a big responsibility in addressing this challenge and should reflect on potential solutions to address these issues within academic circles, and those involving societal stakeholders beyond academia. Indeed, a central social mission of academic institutions is to produce and disseminate reliable scientific knowledge in society. The emerging process through which we try to achieve doing so in our institution (ESCP Business School) has involved three steps. First, there have been calls and ethical charters for “Responsible Use of AI” targeting students and faculty members. These calls are a necessary first step, but they also appear insufficient in and of themselves. Instead, these calls need to be embedded in a broader pedagogical program seeking to alert external stakeholders, in particular students, to the risk of fake news and academic deepfakes. For example, in collaboration with colleagues from ESSEC and HEC, 40 ESCP students (plus 80 from the two other institutions, so 120 students in total) engaged in a two-day seminar during which they came up with solutions to address problems related to fake news. In the run-up to this seminar, students were prepared with three masterclasses focused on fake news and how the phenomenon is linked to emerging technologies such as AI. Last, initiatives are put in place to develop a sensitivity of what good research means in other disciplines, most notably through a conference cycle co-organized with the French

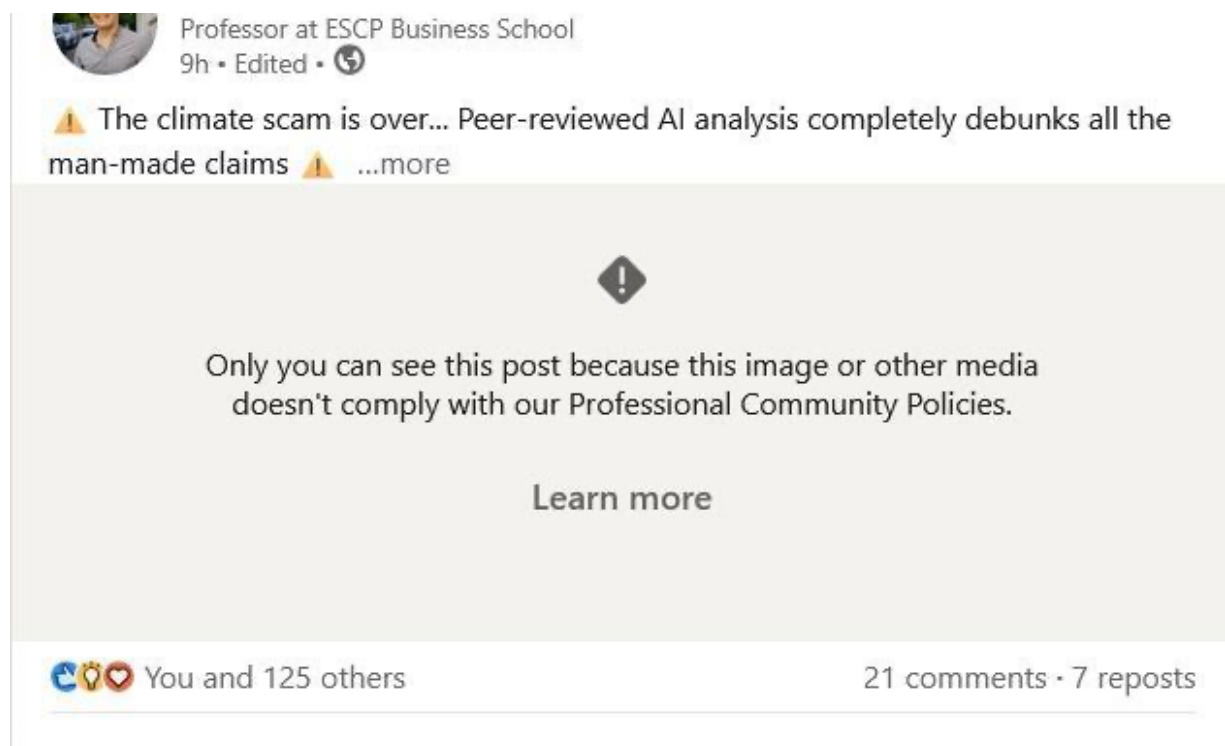
Academy of Science and open to all incoming Master in Management students. Yet, important challenges remain, as AI redefines our relationship to knowledge and truth, and continuous rethinking, sensitivity and vigilance are in order to safeguard a collective understanding of what constitutes reliable scientific knowledge.

Conclusion (and epilogue)

By critically examining the recent publication of a pseudo-scientific academic article written with the help of AI, we have pointed to the worrying emergence of ‘academic deepfakes’ and discussed their associated risks in threatening the perceived objectivity and legitimacy of science, its impact in polluting the scientific debate and the threats it represent for academic impact on society (for example by shaping science-based public policies).

Although extreme, the case reported here shows, in a spectacular way that, contrary to a widely held illusion of AI as a rational, exhaustive and objective tool (Messerli & Crockett 2024), Gen AI technologies are anything but neutral. Because of their opacity and agnostic position towards truth, Gen AI tools may generate significant epistemic risks, and could have deep and unanticipated consequences on academic practice and science.

Epilogue: the recent strange and unpleasant experience of one of the authors suggests another great reason for academics NOT to engage in the arena of wicked debates about academic deepfakes. A mere nine hours after the publication of a LinkedIn post about the deepfake article discussed above, the post was suddenly put offline (without notice) by LinkedIn teams, with no explanation about the reasons other than “this image or other media doesn't comply with our Professional Community Policies” (see image below).



The author discovered, on this occasion, that trying to get information about this issue, or making an appeal is impossible with a free LinkedIn account (unless your account gets banned, which was not the case here as only this post was put offline). After switching to a Premium Account, an online discussion led to the creation of a ticket instructed by the “LinkedIn Safety Team” (with whom it was impossible to interact). After 24 hours, the post was back online, and the author received the following notification email:

"Hi [...],
I hope this message finds you well.
I apologize for any inconvenience.
After reviewing your content again, we confirmed that it complies with our
Professional Community Policies at: <https://lnkd.in/eNe4kpgS>
Your content is now back on LinkedIn.
Thanks again for being part of the LinkedIn community and for helping us
to get our decision right."

No information was provided regarding the reasons the post was momentarily put offline. As the case is closed, no explanation was given by LinkedIn on the reasons for this episode (the main hypothesis being a) automated AI blocking or b) individuals signaling the post to LinkedIn teams). This illustrates again, but in a different way, a key risk with algorithms and AI: their in-built opacity and lack of accountability (O'Neil, 2017).

References

- Beall J. (2017) What I learned from predatory publishers. *Biochemical Medicine* (Zagreb) ; 27 : 273-278
- Chadha, A., Kumar, V., Kashyap, S., Gupta, M. (2021). Deepfake: An Overview. In: Singh, P.K., Wierchoń, S.T., Tanwar, S., Ganzha, M., Rodrigues, J.J.P.C. (eds) *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security. Lecture Notes in Networks and Systems*, vol 203. Springer, Singapore.
https://doi.org/10.1007/978-981-16-0733-2_39
- Chesney & Citron (2019), "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security", *California Law Review*.
- Grok 3 beta, Cohler J., Legates D., Soon F., Soon W. (2025) A Critical Reassessment of the Anthropogenic CO₂-Global Warming Hypothesis : Empirical Evidence Contradicts IPCC Models and Solar Forcing Assumption, *The Science of Climate Change*, vol 5(1) pp 1-16
- 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>
- O'Neil, Cathy. 2017. *Weapons of Math Destruction*. Harlow, England: Penguin Books.
- Oreskes, N., & Conway, E. M. (2010). "Merchants of doubt: how a handful of scientists obscured the truth on issues from tobacco smoke to global warming." 1st U.S. ed. Bloomsbury Press.