WHEN THE BOT ATE YOUR HOMEWORK: DATA MANAGEMENT HABITS WE CAN NO LONGER IGNORE IN THE AGE OF AI

Sorin M.S. Krammer (University of Southampton, UK)

A senior professor toggled a single privacy switch in ChatGPT. In an instant, his two years of carefully structured grant applications, lecture notes, publication drafts, and exam analyses disappeared. No warning. No undo option. Just a blank page. The story, published as a career column in *Nature* in January 2026 *(1)*, attracted widespread sympathy; and understandably so. Losing months or years of precious work is genuinely distressing. But the sympathy, however well-deserved, should not obscure a more important lesson about how individuals and businesses manage their digital work – namelt that over-reliance on commercial AI platforms carries serious, foreseeable risks. Well-intentioned as it is, this account inadvertently normalizes deeply problematic practices among individuals in both academia and business; precisely the audience most likely to adopt its implicit lessons without scrutiny.

A contradiction worth examining

The author acknowledges understanding that ChatGPT produces "seemingly confident but sometimes incorrect statements", yet deemed it reliable enough to serve as the sole repository for two years' worth of academic work. This contradiction is striking. If you don't trust a system's core functions, why trust it as your primary data storage? Further, this concern deepens upon considering that data is the main input for LLM training (meaning that AI companies' incentives are to harness and use as much as possible), and that there are numerous and well-documented concerns about these companies' handling of user data, both ethically and legally *(2)*.

Legal, institutional and ethical dimensions

Treating ChatGPT Plus as a primary repository for academic content is inappropriate for many reasons. First, ChatGPT Plus is not GDPR-compliant *(3)*, a non-trivial concern for any researcher operating within the European regulatory context, for whom this is not a matter of preference but of legal obligation. Second, exclusive reliance on ChatGPT violates one of the basic principles of data management: having multiple backups to avoid critical data loss. Most research institutions have explicit data management policies, many of which have been recently updated to address AI tools usage as well; typically, these policies require that primary research materials be stored on institutional or otherwise secure and recoverable systems. Sensitive

research or academic records should never be deposited on third-party, unvetted platforms. Third, given the scope of this account, we can conclude that ChatGPT Plus was an integral part in the creation of these lectures, grant applications and research publications. This casts doubt on the independence of these outputs. When intellectual work is co-developed iteratively inside a proprietary commercial platform, questions of ownership, originality, reproducibility, and individual contributions become difficult to resolve. These are not peripheral concerns; they bear directly on research integrity, as illustrated by a growing number of cases that span research, teaching, and funding.

Platform accountability versus user responsibility

In the original piece, the professor frames this incident as a failure of OpenAI, flagging that there was no warning or option to undo the deletion. Yet OpenAI did nothing wrong. The company implemented exactly what the privacy-by-design requires: immediate and permanent deletion upon user request. The author even acknowledges this, yet positions himself as a victim rather than taking responsibility for his own lack of judgement and poor digital hygiene. OpenAI's response was not a malfunction; it was the system working precisely as documented and agreed to by the user prior to use. Subsequently, there is no basis (legal or practical) for treating this incident as a platform failure, whether viewed through the lens of responsible AI use or basic common sense judgement.

What responsible AI use looks like

The broader takeaway is not that AI tools are unsafe for academic and professional use- but that integrating them responsibly requires the same basic hygiene we apply to any system. Several principles deserve wider attention:

- Commercial AI chat platforms should never serve as primary repositories for academic work. Local, institutional, or version-controlled backups should always be maintained.
- GDPR compliance and institutional data governance policies must be consulted before using any AI tool with sensitive or confidential research materials. Beyond data loss, deviations from established institutionalized practices may have serious legal implications. Most universities have now published guidance on exactly this.
- Privately owned cloud services are governed by various user agreements that often detail data deletion and training use. Researchers that choose such commercial services bear the responsibility of reading and agreeing with their terms. Reliability and data

stewardship vary considerably across providers, and a monthly subscription fee is no guarantee of either.

In sum, commercial platforms should not be used as primary repositories for academic work, and user error cannot be reframed as platform failure. Clearer interface warnings from OpenAI would certainly have helped, and the experience does point to a genuine design gap worth addressing. But the real cautionary tale here is not about AI tools; it is about the fundamental importance of proper data management, regardless of which systems you use.

AI has a legitimate and increasingly important role in academic and business workflows. But that role requires deliberate governance, not just enthusiasm. The question is not whether to use these tools, it is whether we are prepared to use them wisely. Your bot of choice is not a filing cabinet. Treat it accordingly.

References

(1) Bucher, M. When two years of academic work vanished with a single click. Nature (2026). https://doi.org/10.1038/d41586-025-04064-7

(2) Wikipedia. Artificial intelligence and copyright. *Wikipedia, The Free Encyclopedia.* https://en.wikipedia.org/wiki/Artificial_intelligence_and_copyright (accessed 25 Jan 2026).

(3) OpenAI. Security and privacy. https://openai.com/security-and-privacy/ (accessed 25 Jan 2026).