

## **Bridges Without Foundation? Why the Use of AI Tools in Academia Needs to Build on Ethics First**

AMREI BAHHR, Department of Philosophy, University of Stuttgart, Germany

BAHR, A.: Bridges Without Foundation? Why the Use of AI Tools in Academia Needs to Build on Ethics First  
FILOZOFIA, 79, 2024, No 5, pp. 486 – 500

The paper makes the case for establishing ethically substantiated foundations for using AI tools before turning to more detailed issues. First, I point out AI tools' potential benefits in academia. Second, I discuss the risks of missing out on building ethically substantiated foundations by shifting the focus too much on specific questions. To illustrate this risk, I examine three kinds of issues: problematic outputs of AI tools; the amplification of warped incentives already present in academia; and the risk that universities, other educational institutions, and their members potentially jeopardize their digital autonomy. Third, I present starting points for a discussion on building ethical foundations for using AI tools in academia: we should concentrate on the questions of how to dismantle general warped incentives, what needs to be done to empower institutions in academia to create independent AI infrastructures collaboratively, and how to secure a responsible and productive use of AI technologies.

**Keywords:** AI ethics – responsible use of AI – digital autonomy – digital dependency – open source – warped incentives – research ethics – ethics of higher education

In academia, as in many other domains of our society, AI tools are becoming increasingly ubiquitous – and the same holds for discussions focusing on their use. However, these discussions often center on specific questions: e.g., which limits should apply for using Large Language Models (LLMs) like ChatGPT in exams (Nehlsen – Fleck 2023), whether they should be used as supporting tools for writing grant applications (Moller-Nielsen 2023) or to evaluate applications for professor positions (Hermann 2024). At the same time, fundamental ethical questions get less attention than they deserve. This is especially problematic given that we risk provoking general reservations

towards AI tools as well as over-regulation as excessive reactions if we do not invest in ethical foundations first. This paper aims to provide a starting point to shift the focus of debating AI to these fundamental ethical questions.

### **I. AI Tools: Bridges for Interchanging Knowledge**

Using AI tools in academia promises many benefits, not only for educational purposes but also for research. As the metaphor in this paper's title suggests, such tools can serve as bridges to transport knowledge. This holds not only for propositional knowledge but also for other knowledge forms such as knowing-how: e.g., LLMs can serve to impart the capabilities of structuring an academic paper, assessing arguments, etc. Even critical thinking capabilities can supposedly be taught by so-called "AI tutors" (Benvenuti et al., 2023).

Notably, learning new skills is not only a potential benefit for students but also for researchers who can thereby broaden their skill set, e.g., when they want to adapt to new research topics. Given the enormous databases that AI tools can tackle, the bodies of knowledge out of which relevant items can be picked out and transported, surpass what can be processed by human agents on a large scale. As helpful as it might be, this feature of AI use also notoriously creates certain risks and problems that will be discussed in section II. A of this paper.

The metaphor of the bridge also illustrates that this transfer of knowledge is not happening on a one-way street: users of AI tools in academia are not only recipients of the content AI tools produce. They also provide training data themselves. Certain quality criteria apply to data collected and used in the sciences and humanities to make it suitable for research and education. Hence, academic users will likely provide AI tools with training data of exceptionally high quality compared to data collected in other areas that do not operate with similarly strict quality criteria. The same should be valid for many of the prompts. Often, these prompts will result from pre-structuring by researchers competent for the topic and hence also be informed by relevant insights and findings from pertinent research.

Moreover, given that academic users should be familiar with the principles of good scientific practice<sup>1</sup>, we can at least hope that their use of AI tools happens against the backdrop of these principles. Hence, the inputs provided by users from academic contexts can be valuable training data for

---

<sup>1</sup> The German Research Foundation (DFG) provides a specification of these principles, see Deutsche Forschungsgemeinschaft (2022).

the AI tools being used. We should keep this aspect in mind, as it will be crucial for the question of digital autonomy discussed in section II. C.

## **II. When the Use of the Bridges Casts Doubt on Their Reliability: Application of AI Tools in Academia and Its Multiple Risks**

Using AI tools in academia and other educational contexts can be beneficial in several regards. However, the potential risks of this use are neither less diverse nor significantly less likely to come about. To give an outline of the different kinds of risks at play here, I will distinguish between risks pertaining to problematic inputs and outputs (section A), risks that are provoked by making use of AI tools against the backdrop of certain warped incentives already at play in academia (section B), and risks affecting the conditions of responsible AI use, namely by compromising the digital autonomy of individual and institutional academic agents (section C).

### **A. Fake News, Discrimination, Plagiarism, and Hallucinations: Questioning the Bridges Because They Serve to Transport Problematic Content**

Despite all the benefits that AI tools can have, they can also serve to transport content that is less beneficial or even detrimental for academic purposes. This includes the risk of additional collateral damage if problematic results produced here are applied to other areas (e.g., if research findings are used in policy advice). Given that AI tools draw upon a vast set of heterogeneous content as training material, the details of which often remain opaque, it is not trivial to ensure that the tools do not produce problematic content.

Among the problematic outputs of AI tools that have prominently been discussed so far are, e.g., the reproduction of content that is discriminating or intentionally untruthful, outputs that violate authors' rights by making use of their work without their consent, and so-called AI hallucinations: as we know, LLMs such as ChatGPT sometimes make up scientific literature that does not exist, and even present us with elaborated back-stories to such literature upon request. The possibility of such hallucinations poses a massive problem for using AI tools in academia.

However, hallucinations are not the only phony outputs occasionally produced by AI tools. Given the vast amounts of data they are trained with, the reproduction of fake news is another risk we face. This holds not only for

outputs in the form of text but also for images and videos generated by AI tools (Stöcker 2024).

The reproduction of biases, discrimination, and injustice is another massive problem already being discussed in the literature (e.g., Cossette-Lefebvre – Maclure 2022). It should be indisputable that responsible use of AI technology (not only) in academia requires precluding such problematic features (even though the question remains which instances of AI outputs exactly are to be subsumed under these problematic categories).

Another issue is the violation of authors' rights by using AI tools, the results of which are regularly described as plagiarism in recent discussions.<sup>2</sup> That the topic of plagiarism produced by AI tools prompts heated discussions outside academia as well is prominently illustrated by the lawsuit that the New York Times issued against OpenAI and Microsoft (Grynbaum – Mac 2023): the New York Times sued OpenAI and Microsoft, claiming that millions of *The New York Times* articles have been used for training the LLMs of these companies. AI tools use giant sets of diverse content for training, some of which contain material to which authors' rights apply while not revealing in detail which content they have processed. Hence, it is no surprise that the risk of plagiarism also needs to be considered concerning the academic use of AI, not least because of the rules of good scientific practice.

One might object that the risks mentioned above can be dealt with by thoroughly examining what AI tools have produced. However, this would require checking every element of these outputs, especially when applied in areas outside the user's expertise – or dispensing with the application of these tools altogether when it comes to such areas. Both would significantly limit the benefits of using AI tools, so we should treat this as a last resort.

One might also object that further development of AI tools should help to detect the kinds of problematic outputs mentioned above and ultimately prevent them from being produced in the first place. However, even if both this and an evaluation and improvement of the material that is being used to train AI tools might very well be possible in the long run, problematic outputs

---

<sup>2</sup> Notably, given that we do not know how their output comes about, distinguishing between hallucinations and plagiarism is not trivial: a text that is notationally identical to an already existing text might be the result of copying and hence count as plagiarism, but it could also be the result of hallucinating AI, which would make it a parallel creation that exemplifies notational identity only coincidentally. Depending on which of the two categories a text belongs to, different ways of dealing with it are required in academic contexts (see Bahr 2024 for details).

of these kinds will nonetheless remain one of the main risks of using AI tools at least in the near future.

Moreover, another problem arises concerning potential upgrades of AI tools if we refrain from using AI tools altogether until such upgrades are available. Not only are we depriving ourselves of their benefits, but we might also even contribute to hindering their improvement. Refraining in principle from using AI tools because they are not reliable enough yet will likely stabilize their inadequacy. Improving these tools requires their very use to ensure that they “learn” to get better, also concerning risk-prevention. Here, parallels could be drawn to the debate about providing personal health data: to make sure that such data can be used for the benefit of the public, individuals need to be willing to provide their data to ensure that sufficient data can be collected to, e.g., learn more about certain diseases. In this context, it has been discussed whether there is a duty to share such data (and as a part of this debate, it has been argued that using data without explicit subject consent can be viable at least in some instances, see, e.g., Ballantyne – Schaefer 2018). However, a similar argument for providing data by using AI tools should be treated with caution: it is precisely our contribution to the improvement of certain AI tools that should raise concerns regarding our use of these tools. This holds at least as long as they are in the hands of tech companies, but also concerning their pending indispensability, as will become apparent in section C.

Regarding the context of academia, additional risks to the ones just mentioned need to be considered. While the risks discussed above apply equally to other societal domains, specific risks are imposed on using AI tools in academic contexts. Research on the use of LLMs in educational contexts points to worrying consequences here: distortions and biases in the training data and algorithms on which these models are based threaten to lead to problematic normative monocultures, narrow perspectives, or one-sided representation of content (Bahr – Mayer 2024). Digital infrastructures shape both the type of knowledge that researchers can generate and how they interact with each other in their daily practice (see, e.g., Kaltenbrunner 2017). This calls for diligence when setting up infrastructure for academic contexts relying on AI technology.

In addition, as long as the bridges built with the help of AI are regularly used to transport problematic content, making use of them in academia, despite these risks, can also play into the hands of those questioning academic endeavors in science and the humanities altogether. In recent times, we have

seen a rise of alarming tendencies such as growing skepticism, which has even culminated in hostility towards science and the humanities. As long as we lack reliable and realistic means to ensure that the outputs of AI tools remain free from problematic constituents of the kinds just mentioned, using them in academic contexts involves the danger of fueling such worrying developments even further. Hence, upholding the standards of good scientific practice, no less when using AI, is crucial to rebuilding and strengthening societal trust in academia, its agents, and its results. For the sake of trust in scholarly work and education, we need to build ethically reliable foundations for using AI in academia from the start.

### **B. Warped Incentives in Academia Impend to Be Fueled Even Further by the Use of AI**

Even if the risks mentioned in section A could be effectively contained with optimized AI tools, a cautious approach is nonetheless preferable: one that first questions the inflationarily growing, partly misguided demands in the academic system as such instead of simply continuing to serve them with AI support.<sup>3</sup> After all, it is not a necessary condition for a functioning academic system that researchers are constantly beating themselves up to meet a myriad of requirements, most of which are quantitative. For instance, the principle *publish or perish* is obviously not a value in itself. Instead, it is a warped incentive that produces unwanted consequences. The number of publications is already so large that we must ask ourselves who can – and should – read publications in this excessive quantity. The fact that the number of publications can now even be increased somewhat effortlessly using chatbots and similar tools is, without question, not a particularly promising scenario (Bahr 2024).

The same applies to third-party funding, which has long since taken on questionable proportions: as promising as it may sound to use AI tools for writing and even reviewing grant applications, this would likely accelerate undesired developments in funding like the overemphasis on grant money and competitive procedures even further. Instead of satisfying the academic system's hunger for grant money, it could make the system even more insatiable in this regard. Our options to use AI tools should instead serve as a reason to question this system in the first place (Parrilla 2023).

As long as such warped incentives in the academic publication system, the system of research funding, etc., remain active, however, using AI tools

---

<sup>3</sup> See Rahal et al. (2023) for an overview of these problematic demands, their harmful effects, and suggestions for overcoming them.

might worsen them even more and thereby intensify the harmful effects on the individuals involved, the academic system as a whole, and its scientific outputs.

### **C. Loss of Digital Autonomy Through AI Use in Academia**

Another enormous risk presented by a lack of ethically thought-through foundations for AI use in academia is the loss of digital autonomy. Here, we can roughly understand digital autonomy as the capacity to perform independent, self-determined actions with regard to our use and disuse of digital tools and other digital entities – in what follows, its features should become clearer once we consider what puts it at risk. So far, the topic of digital autonomy and the relevant factors for its potential loss have mainly been discussed regarding agents outside the scope of academia, e.g., countries in the EU (Mayer – Lu, 2023). However, it is no less crucial in our context. The effects of AI tools on the organization of research and teaching in academia are already visible – not only in the sense discussed in section A but also with regard to shaping the very conditions of research and teaching. Digital infrastructures significantly influence the social and epistemological organization of research fields and practices as well as those of teaching and learning already (Kaltenbrunner 2017), and these effects will most probably increase even further by using AI tools in the near future. At the same time, agents in academia are notoriously lagging behind not only in trying to understand the rapid changes of their digital environments (and, as a result, also the environments of research and teaching that these are part of) but also in actively shaping these environments. Academic agents are left far behind by the tech companies involved, which already dominate the discourse about socio-technical designs for the future of academia and our education systems, learning, and the co-production of knowledge (Mager – Katzenbach 2021).

One worrying result of these developments is that agents in academia – universities, researchers, lecturers, or students – run the risk of losing their digital autonomy altogether. Even if we leave aside the relevant argument here that many of the cases at hand cannot reasonably be described as cases of autonomous decision-making in the first place because users lack relevant information to give their informed consent<sup>4</sup>, another aspect is even more

---

<sup>4</sup> A similar argument has been made regarding notice-and-consent practices, even though it might be possible to save their normative power regardless (Susser 2019).

worrying here: choosing to use certain AI tools, though being an autonomous decision in itself, can ultimately result in a substantial loss of digital autonomy.

This might sound puzzling or even paradoxical at first: did the actors involved not make an autonomous decision to use these tools to begin with? How can such an autonomous decision endanger autonomy as such? To substantiate the claim that the decisions for using AI tools are autonomous on the level of the involved institutions, it could be argued that universities do indeed decide autonomously to use external digital tools and even base these decisions on established democratic processes in university committees. So, how can such autonomous decision-making ultimately result in a loss of autonomy?

To see why this could be the case, we should note that autonomous decisions reach their limit when they undermine autonomy as such. This has been discussed regarding personal autonomy, e.g., in the realm of sexuality. In the context of BDSM practices, Nielsen argues that

consenting to something which undermines [someone's] opportunity of exercising autonomy (and/or something that will undermine the very capacity for autonomy in the long run) is void (Nielsen 2010, 273).

Even if the requirements for autonomy are fulfilled regarding the decision as such, this decision ultimately leads to a loss of autonomy:

If B – being at the time of consent a free, uncoerced, rational, person with a reasonably clear picture of the consequences of his or her actions and being sufficiently clear and stalwart in his or her intentions – consents to a contract where he or she is to be held captured and disciplined into blind obedience for an indefinite period, then B in a very real sense ‘consents’ to having his or her autonomy eradicated (Nielsen 2010, 272).

To analyze in how far this argument can be applied to our topic, I will distinguish four kinds of dependencies and discuss whether they endanger autonomy in a similarly profound manner as in the BDSM example. I will indicate who the dependent subject – individual and/or institution – is for each kind of dependency. Also, I will show how these subjects are related with regard to the respective kind of dependency. Notably, it is not immediately clear whether the BDSM argument, which centers around personal autonomy, directly translates to cases of institutional autonomy. However, regarding the



individual agents in academia, this translation should present less of a problem, so I first turn to these agents.

Suppose individual agents willingly tie themselves into using certain AI tools in a way that cannot easily be revoked. For instance, if individual academic users extensively rely on AI tools in their research, teaching, or learning, they might run the risk of making themselves dependent on their use in the long run, e.g., by relying on them to develop research questions or to write well-structured research papers, instead of acquiring and improving the skills to do these things on their own. We can call this a *dependency due to deskilling*. Note that this problem is not exclusive to commercial AI tools: the solutions provided by universities are also not immune to creating this kind of dependency. Here, the dependent subjects are the individual researchers and students. Nevertheless, individual dependency transfers to institutions once a significant number of their members cannot attend to their tasks without the help of AI tools anymore. Nonetheless, this kind of dependency does not pose an analogous problem as in the BDSM example: here, autonomy is eradicated only in part, temporarily and reversibly, given that the lacking skills could still be acquired.

Despite not being too serious anyway, preventing dependency due to deskilling is more challenging than it seems. An individual decision to opt out of using AI tools in academia to preserve one's skills might soon be near impossible due to another impending dependency: once most people resort to AI tools, opting out of their use would result in massive drawbacks. Considering the competitive logic in academia, refraining from using these tools might ultimately result in individual competitive disadvantages, e.g., because fulfilling tasks without AI takes much more time. I suggest calling this a *competition dependency*.

Subject to this dependency are not only researchers and students but also institutions: if all or most universities establish the use of AI tools, e.g., to keep up with the competition presented by university rankings, refraining from their use might also become untenable for universities. As long as the current competitive logic of academia remains untouched, this dependency cannot be revoked easily – though it could be in principle, namely, if the academic community were to change this logic. So, even though not irreversible, this dependency cannot easily be untied.

This is all the truer for another, much more severe component of competition dependency: commercial AI tools benefit massively from being used in academia because they improve through training data. Given the high

quality of the data provided by academic users, these tools will become even more powerful the more they are used in academia. Hence, by using these commercial tools, agents in academia – both on the individual and institutional level – contribute to the fact that they can no longer keep up with the competition from tech companies once they come up with their own ethically founded AI solutions. Combined with the other component of competition dependency just mentioned, this is comparable to the BDSM example: if universities want to keep up with the competition, they are tied to commercial tools, which have only become indispensable through their own, initially autonomous use.

In addition, by choosing to license AI tools provided by tech companies, universities are tying themselves to those companies in a way that might be difficult to unravel again. This is a kind of *contractual dependency*, which, though not permanent in principle, is binding at least for a certain period. Hence, by entering into such contracts, the autonomy of academic institutions may at least be compromised temporarily.

Above, I have pointed out that the evolution of AI tools might be hindered if individuals decide not to use them and thereby withhold input that could improve them (see the health data analogy in section A). Given that powerful AI tools could serve the common good, it could be argued analogously that institutions are responsible for supporting the evolution of these tools by providing training data. Suppose they withdraw their use of AI tools (and thereby also the use by their members). Would this not hinder the development of such tools to a significant degree, given that sufficient amounts of high-quality data are required to train those tools properly? We can call this a *dependency on behalf of the common good*, which applies to individuals and institutions.

However, we should be careful not to give in to this argument too carelessly. After all, AI tools provided by tech companies primarily serve the purpose of profit rather than that of the common good. Using these commercial tools might even lead to a substantial loss of digital autonomy on both the individual and institutional levels, as we will see in what follows. Nevertheless, things might turn out differently for non-profit open-source solutions.

To sum up, competition dependency is the kind of dependency that resembles the BDSM case the most: the decision to use commercial AI tools can eventually result in eroding the institution's digital autonomy altogether, because by improving commercial tools through their use, these tools will ultimately be so powerful that universities cannot catch up and develop

comparable own solutions anymore. Even though dependency due to deskilling, dependency on behalf of the common good and contractual dependency are less severe, we should be aware that they can compromise digital autonomy at least partly and/or temporarily.

### **III. Doing the Groundwork First: How to Build Reliable Bridges with AI in Academia**

The aforementioned risks inflicted by using AI tools in academia call for developing ethical foundations. Such foundations can ensure that AI tools serve as reliable bridges for the transportation of knowledge, which are beneficial for academic endeavors instead of harming them. Note that this is not only required from an ethical perspective but there is also a pragmatic argument for creating such foundations. Suppose we proceed with using AI tools as we currently do without building such foundations first. In that case, we risk provoking overcompensation, whether in the form of general skepticism regarding the use of AI tools or its overregulation. In what follows, I will therefore sketch two main leverage points for building such foundations: the creation of independent digital infrastructure that can serve as a groundwork for AI use and the substitution of warped incentives by more suitable basic conditions in academia to make sure that AI use does indeed help and not hinder productive academic endeavors.

#### **A. Developing Independent AI Infrastructure for Academia Instead of Relying on Unpredictable Tech Company Solutions**

A crucial leverage point for improving the use of AI tools in academia is the self-development of independent structures by academic institutions. This should happen by teaming up and creating shareable, non-profit, open-access, and open-source digital infrastructure for academic purposes (Bahr – Mayer 2024). Note that this does not mean that, e.g., LLMs need to be created by universities themselves. Paaßen (2024) points out that they can instead resort to existing open-source models, which are available, e.g., on *huggingface.co*. According to Paaßen, in addition to these models, universities need two other elements to create their own AI solutions:

Secondly, the computing infrastructure to run the models, especially servers with powerful graphics cards. Thirdly, and most importantly, the expertise to select the most powerful LLMs currently available, install them

on the university's infrastructure, and provide easy-to-use interfaces for learners and teachers (Paaßen 2024, my translation).

As Paaßen points out, universities are exceptionally qualified regarding the third element, given their research on AI and LLMs.

The resulting university-based solutions could also help tackle the problems of discrimination and bias, at least in part, given that one source for these unwanted features of AI outputs is what Cossette-Lefebvre and Maclure (2022) call the opaque nature of the respective tools: we do not know how exactly the results of those tools come about. This is increased even further if the tools used are in the hands of tech companies that do not disclose their operating principles. Hosting and maintaining their own tools could help academic institutions limit this problematic opacity to a minimal degree. Given that opacity is one source of reproducing biases and discrimination, adding transparency and accountability in this way could be a promising antidote to these problems, at least in academic contexts.

Notably, establishing the respective infrastructure in an ethically convincing way is a challenging undertaking. Several questions remain to be answered. For instance, how should we deal with the trade-off between limiting the training data to content that does not affect authors' or personality rights, resulting in less powerful AI tools, and the creation of more powerful tools that can keep up with the tools provided by tech companies (which regularly do not strictly impose such ethical considerations on their tool development)? Developing, mapping, and answering the questions required to create independent, ethically founded AI solutions for academic purposes should be a task for future research on the use of AI in academia.

## **B. Establishing New Basic Conditions in Academia to Build Stable Bridges with AI**

As for the warped incentives potentially amplified by AI use, we should prioritize their abolition first instead of praising AI tools as a potential means to feed them even further. To begin with, as a scientific community, we should ask ourselves whom we are serving if we continue to maintain the excessive requirements imposed on us as researchers. Science? Ourselves? Both are highly questionable. Most likely, neither researchers nor research itself benefits from such a large number of scientific publications that can barely be acknowledged by anyone. Those who do benefit from the system, however, are the publishers, who make money from publications and hence prefer a vast

number of them. But this should hardly be a reason to adhere to the requirements in question (on the contrary, there are good reasons for us to take a critical stance on the business models of some scientific publishers anyway).

AI tools can bring massive changes to academia and have several beneficial potentials. But it is up to us to shape this process in a goal-oriented way by first eliminating warped incentives and stopping undesirable developments, above all, the problematic tendency towards quantity as a quality feature. In an academic system that rests on competitive procedures focusing on quantitative measures, chronic use of tools that can serve to produce even more papers, grant applications, etc. poses the risk of perverting the exaggerated standards even more. Hence, instead of praising those tools for their potential, we should instead come up with alternative standards of evaluation that shift the focus from quantity to quality once again, e.g., as suggested by CoARA. Once this has been done, the second step should be to ask ourselves what AI can do to improve our academic endeavors, whether in research or teaching. If we want AI tools to change work in academia for the better, we should proceed in this order instead of further exacerbating existing misguided incentives and developments through their use. Hence, it should be worth taking up insights gained in the discussions about reputation mechanisms and key figures in academia and connecting them also with questions regarding the use of AI tools in future research.

#### **IV. Conclusion**

Undoubtedly, using AI tools in academia can be beneficial in several respects. However, as we have seen, different kinds of risks of such use give cause for concern: the (re)production of fake news, discrimination and bias; the production of plagiarism and hallucinations; the intensification of detrimental effects produced by warped incentives at play; and last but not most minor, the risk of forfeiting digital autonomy altogether. This raises the question of how the benefits of AI use can be reaped in academic contexts while simultaneously limiting its risks as much as possible. Along these lines, I have presented two leverage points to build effective, ethically reliable foundations for the use of AI tools: the creation and deployment of own solutions by academic institutions or coalitions of these institutions, respectively, and the dismantling of problematic conditions that shape academia in an undesirable manner anyhow, and whose negative consequences run the risk of being

amplified even more if AI tools enter into the equation before these conditions have been abolished.

## Bibliography

- BAHR, A. (2024): Same same but different: Textidentität, Autor\_innenschaft und Schöpfungshöhe im Angesicht von generativer KI. In: Schreiber, G. – Ohly, L. (eds.): *KI: Text. Diskurse über KI-Textgeneratoren*. Berlin – Boston: De Gruyter, 169 – 182. DOI: <https://doi.org/10.1515/9783111351490-012>
- BAHR, A. – MAYER, M. (2024, January 15): Auswege aus der digitalen Unmündigkeit. *Frankfurter Allgemeine Zeitung*, 4.
- BALLANTYNE, A. – SCHAEFER, G. O. (2018): Consent and the ethical duty to participate in health data research. *Journal of Medical Ethics*, 44(6), 392 – 396. DOI: <https://doi.org/10.1136/medethics-2017-104550>
- BENVENUTI, M. – CANGELOSI, A. – WEINBERGER, A., et al. (2023): Artificial Intelligence and Human behavioral Development: A Perspective on New Skills and Competences Acquisition for the Educational Context. *Computers in Human Behavior*, 148. DOI: <https://doi.org/10.1016/j.chb.2023.107903>
- COSETTE-LEFEBRE, H., – MACLURE, J. (2023): AI's Fairness Problem: Understanding Wrongful Discrimination in the Context of Automated Decision-Making. *AI Ethics*, 3, 1255 – 1269. DOI: <https://doi.org/10.1007/s43681-022-00233-w>
- Deutsche Forschungsgemeinschaft. (2022): *Guidelines for Safeguarding Good Research Practice: Code of Conduct*. Bonn: Deutsche Forschungsgemeinschaft. DOI: <https://doi.org/10.5281/zenodo.6472827>
- GRYNBAUM, M. M. – MAC, R. (2023): The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work. [Online.] *The New York Times*. Available at: <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>
- HOLMES, W. – TUOMI, I. (2022): State of the Art and Practice in AI in Education. *European Journal of Education*, 57, 542 – 570. DOI: <https://doi.org/10.1111/ejed.12533>
- KALTENBRUNNER, W. (2017): Digital Infrastructure for the Humanities in Europe and the US: Governing Scholarship through Coordinated Tool Development. *Computer Supported Cooperative Work*, 26, 275 – 308. DOI: <https://doi.org/10.1007/s10606-017-9272-2>
- MAGER, A. – KATZENBACH, C. (2021): Future Imaginaries in the Making and Governing of Digital Technology: Multiple, Contested, Commodified. *New Media & Society*, 23 (2), 223 – 236. DOI: <https://doi.org/10.1177/1461444820929321>
- MAYER, M. – LU, Y.-C. (2023): Illusionen der Autonomie? Europas Position in den globalen digitalen Abhängigkeitsstrukturen. *SIRIUS – Zeitschrift für Strategische Analysen*, 7 (4), 390 – 410.
- MOLLER-NIELSEN, T. (2023, December 19): European Research Council Issues Warning on AI's Use in Grant Applications. [Online.] *Science Business*. Available at: <https://sciencebusiness.net/news/ai/european-research-council-issues-warning-ai-use-grant-applications>

- NEHLSSEN, J. – FLECK, T. (2023, March 31): Ist ChatGTP ein zulässiges Hilfsmittel in Prüfungen? [Online.] *Forschung & Lehre*. Available at: <https://www.forschung-und-lehre.de/recht/ist-chatgtp-ein-zulaessiges-hilfsmittel-in-pruefungen-5524>
- NIELSEN, M. E. J. (2010): Safe, Sane, and Consensual: Consent and the Ethics of BDSM. *International Journal of Applied Philosophy*, 24 (2), 265 – 288. DOI: <https://doi.org/10.5840/ijap201024223>
- PAASSEN, B. (2024, January 26): Warum Hochschulen jetzt eigene Sprachmodelle hosten sollten. *Jan-Martin Wiarda Blog*. [Online.] Available at: <https://www.jmwiarda.de/2024/01/26/warum-hochschulen-jetzt-selbst-sprachmodelle-hosten-sollten/>
- PARRILLA, J. M. (2023): ChatGPT Use Shows that the Grant-Application System Is Broken. *Nature*, 623 (7986), 443. DOI: <https://doi.org/10.1038/d41586-023-03238-5>
- RAHAL, R. M. – FIEDLER, S. – ADETULA, A., et al. (2023): Quality Research Needs Good Working Conditions. *Nature Human Behaviour*, 7, 164 – 167. DOI: <https://doi.org/10.1038/s41562-022-01508-2>
- STÖCKER, C. (2024): Lernende Maschinen und die Zukunft der Öffentlichkeit. In: Schreiber, G. – Ohly, L. (eds.): *KI: Text. Diskurse über KI-Textgeneratoren*. Berlin – Boston: De Gruyter, 401 – 418. DOI: <https://doi.org/10.1515/9783111351490-025>
- SUSSER, D. (2019): Notice after Notice-and-Consent: Why Privacy Disclosures are Valuable Even if Consent Frameworks Aren't. *Journal of Information Policy*, 9, 37 – 62.

---

Amrei Bahr  
Department of Philosophy  
University of Stuttgart  
Seidenstraße 36  
70174 Stuttgart  
Germany  
e-mail: [amrei.bahr@philo.uni-stuttgart.de](mailto:amrei.bahr@philo.uni-stuttgart.de)  
ORCID ID: <https://orcid.org/00000-0002-9527-2601>