



ELEPHANT

IN THE LAB

INTERVIEW

The (potential) impact of AI on the individual research process and science in general.

Short title	The (potential) impact of AI on the individual research process and science in general.
Long title	The (potential) impact of AI on the individual research process and science in general. An Interview with Theresa Züger
Authors	Theresa Züger ¹ , Sascha Schönig ¹
Author affiliation	¹ Alexander von Humboldt Institute for Internet and Society
Author bios	<p>Theresa Züger is head of the Public Interest AI research group funded by the German Federal Ministry of Education and Research (BMBF). The group is concerned with the question of how AI can serve the public interest and which technical and social criteria must be fulfilled for this to happen. She is also lead of the AI & Society Lab. The Lab was launched by HIIG in 2020 and is an inter- and transdisciplinary research lab at the intersection of academia, industry and civil society. Theresa's research focuses on the political dimensions of digital technologies and cultures, with particular interest in questions of democratic theory.</p> <p>Sascha Schönig is a blog editor for Elephant in the Lab and works at Alexander von Humboldt Institute for Internet and Society in the project Capaz (Capacities and competencies in dealing with hate speech and hostility towards science).</p>
Author social links	Theresa Züger: LinkedIn Website Sascha Schönig: ResearchGate Email
Date published	March 22nd, 2024
DOI	10.5281/zenodo.10850268
Cite as (APA)	Züger, T., & Schönig, S. (2024). The (potential) impact of A.I. on the individual research process and science in general. An Interview with Theresa Züger. <i>Elephant in the Lab</i> . DOI: https://doi.org/10.5281/zenodo.10850268

INTERVIEW

To launch our new section 'AI in Research' on [Elephant In The Lab](#), Sascha Schönig spoke to Theresa Züger, head of the Public Interest AI research group, about the influence of AI on her personal day-to-day work in research, as well as the science system as a whole. She gave some exciting insights about the risks and opportunities AI bears for research work and talked about tools her team is developing at the Alexander von Humboldt Institute for Internet and Society. She advocates for 'a high standard of transparency in AI, not only for open science purposes, but also for the purpose of an ecosystem of public interest AI and research on AI'.

Sascha Schönig: AI has the potential to change research. There's a lot of positive potential, but at the same time, there are also fears associated with the influence on academic activity. In your impression, how enthusiastic are researchers at the moment about future effects of AI?

Theresa Züger: That's hard to say because I don't know all fields of research and I think that different parts of the community will have to expect very diverse types of changes. In the very start of the AI hype researchers of all disciplines had very high hopes; not only the computer sciences, data science or natural science who already have a history of working with AI models, but also researchers from social sciences and the humanities who were asking themselves how AI can help doing their jobs.

Some years later, most researchers have realized that AI will not simply take over responsibilities or do jobs for us as we want them because there are always glitches that we are aware of and there are limits that we need to be aware of. For now, we understand that we can't speak of a general purpose AI that would fit all needs. It also means that it's only going to be solidly good in particular tasks, but there will always be glitches if you look from a very specific angle, and researchers always do. So if researchers are working with AI in a very specific way, and a lot of researchers are, then I think they have been quite aware of the potential, even before ChatGPT, and they're also quite aware of the limitations. That's at least my guess. So in my team, we're also using AI and we're aware of what it can do for us, but we're also well aware when we can't use it for research and where it's not helping.

Sascha Schönig: Yes, it was wonderful because most things you mentioned I would've asked anyway. In the beginning when you talked about scientific fields you mentioned data scientists and computer scientists. Would you say that these are the key areas where AI has progressed the most to this point? So are those the key scientific areas where it can help the most?

Theresa Züger: That's a difficult question, because in the areas you mentioned the advances brought by AI are most clearly visible, if we speak of modeling and achieving certain metrics in order to have certain performances in these models. This is interesting only to a certain point, the bigger question is how can AI be of use in interdisciplinary research and in very specific fields such as in medical or pharmaceutical research, geology, biology or genetics. These are fields where the use of AI will increase. What is important to keep in mind, that AI will influence

INTERVIEW

research that has a direct impact on society, and here the responsibility rises, I mean, in these cases it is extremely important that a model performs well, and not only in regard to a certain metric, but to its actual output.

And for that, models don't always need to be the fastest, most sophisticated or strongest but they need to be specific and very clearly adapted to a task. In this case a very curated data set is needed that relates to a very specific problem that AI can help with. And in this respect, I think the interesting things actually happen in combination with other experts from computer and data science for example as in my team. So you need expertise from different research areas and fields to understand most problems that are of societal relevance. Hopefully, we will see more and more of such interdisciplinary teams coming together.

Sascha Schönig: So, in your opinion, is a lack of interdisciplinary teams a challenge when it comes to using AI for research?

Theresa Züger: It's definitely one challenge, yeah.

Sascha Schönig: Do you have more challenges or limitations in mind because you've mentioned in the beginning that there are many limitations right now regarding AI and research or science in general?

Theresa Züger: Interdisciplinarity is definitely one of the biggest challenges. I've been leading an interdisciplinary team myself for the last couple of years and I learned very much in these years. The chances to produce a technology by engaging computer and data scientists with colleagues from other disciplines like the social sciences helps to achieve a bigger impact and to address a societal problem more to the point. Of course, this is always very context specific.

So, yes, I think we need much more inter- and even transdisciplinary work where more practitioners are involved in the work around AI, specifically when it comes to addressing societal problems where science aims not only to produce theoretical insights but also knowledge that helps addressing societal challenges.

And you probably know that this is the part where science needs to improve - producing knowledge that is particularly relevant for citizens. This kind of work is also not rewarded enough in academia, we can count ourselves lucky in my team that for us it's a big reward to know that we have an impact. But this is rather an exception and doesn't happen so much in the academic system, so I think we need to break the silos of science and stop working on problems in isolation. And partly we see projects where this is already happening, especially in our field - in the work around [Public Interest AI](#) which is really cool.

Sascha Schönig: At the beginning you mentioned that you use AI tools in your daily research. May I ask, which tools you are currently using and for what kind of tasks?

Theresa Züger: Let's start with an obvious tool like ChatGPT. I've tried it out a couple of times and every time I was disappointed as a researcher. For me it's not really useful and I can't think of any tasks that ChatGPT would be suitable for, because I want much more specific answers to my very specific questions. On the other hand, there is a positive side to it; it is a confirmation that my work is important because a tool like ChatGPT cannot just handle my tasks. In our team I know some are using ChatGPT because they're researching it to understand how AI functions but we're actually not using it to produce anything.

What I do sometimes use is [Elicit](#), it is a tool where you can sign up and get actual research papers with a little summaries of the paper to any research question that you have. And this is a really good entry point to get into a topic if it's a field that you're not so experienced in, I would describe it as Google-Scholar-Plus idea. So, you basically get an updated and ready to use research database generated by AI. I guess there probably might be similar tools in the making. In this sense, it seems reasonable to use AI in research just to speed up your process a bit.

However, we are building AI ourselves. So in our team, like the biggest part of how we work with AI is we're training it. We have two projects. One is a browser plugin called [Simba Text Assistant](#), where you can enter German text and you get a simplified version. My colleague Freya Hewett developed it and is writing her PhD on simplified language in German and she trained a data set herself and got one from the Austrian press agency. She annotated the data and trained it herself to tweak other open source models for the task of translating to a simplified text in German. And it's a tool that we hope helps language learners to better understand content that is too complex. It also has a function that underlines complicated words and explains them to the users. We are also hoping to work with different domains, very applied, where they can actually, from their domain knowledge, enter important words and explain them in simpler words. And they will pop up in whatever texts their students will enter. It is basically an open source wiki that explains complicated words using simplified language.

Sascha Schönig: Wow, this is pretty cool because I'm thinking of science communication, when you try to simplify it for an audience which is not a scientific one, then tools like this would be so great to have.

Theresa Züger: Yes, I see numerous potential fields of application and we hope that it takes off in the community. It was mainly Freya who trained the model with some help from co-workers and she produced a browser plugin that is going to be published in the next coming weeks. There is another interesting project in this context where another colleague of mine - Sami Nenno - is working on the issue of disinformation and fact checking. He collaborates with many fact checking agencies that are really overwhelmed by checking potential fakes. This kind of work is very context specific and needs to be done by humans. We are trying to help develop a better filter that could point to pieces of information that need to be checked and we hope that this will make their work more efficient. Actually, filtering and assessing claims takes much more time

than checking the information for fakes and disinformation itself. And he's collaborating with fact checkers like [Correctiv](#) and many others to build something that is useful for them.

We try to have an impact and produce something that can be of use for practitioners in the field; this is the reason why we have training infrastructure ourselves. While doing research in the area of public interest AI it makes a huge difference when we're able not only to talk from a theoretical standpoint, but also know what the actual problems are when you train a model like this yourself.

Sascha Schönig: In your opinion, what are ethical considerations around AI and scientific research? There are a lot of reservations or fears of negative effects AI could have in the future and already has when it comes to training models like biases in the data.

Theresa Züger: There are so many ethical considerations. I think, first of all, every researcher should ask themselves, do we need this in the first place? Because you need to invest a lot of resources if you want to use AI in your work; and you invest them into a system that will not give you certainty about a decision but a probability. It's amazing what this opportunity can do for some research areas, but in some cases there might be better architectures and solutions that would provide higher levels of certainty. So one should not simply follow the hype and start using AI because everyone does, but really look into the details and check whether this is the best architectural solution for a problem or if there is another, more suitable, one. And if one decides to use AI in his or her research, it's important to consider sustainability and develop models that can be reused by others as well as reduce the data to the least possible. Most research is publicly funded, so, in my opinion, the research community should benefit from it.

But there are so many other ethical issues worth mentioning. Depending on what you train, there is always a potential for bias and discrimination. And I think researchers need to be very cautious about this. For instance, in medical research, I think sometimes it is not enough to think about bias mitigation techniques but to really ask yourself the ethical question if considering the risk of discrimination in this field can make things better. These things will be answered very differently in every specific case, but I think we need to have a very high level of scrutiny in our research.

And then, of course, there is the question of quality. Who is annotating the data? Behind every AI project there are a lot of hours of mostly invisible human labor. In our case, we did this annotation together with student assistants, they were properly paid for their work, we did not outsource it to an unknown offshore company in a different country. And I think that researchers have a big responsibility to also make sure that their work is done under proper ethical conditions.

What also needs to be considered are, for instance, questions like: If your projects involve very specific societal groups, how can you include them? How can they have a voice in the data governance or in the governance of the tool?

INTERVIEW

So there're many not only ethical but also societal and political questions attached to this. I think that everybody working on AI projects in any societal field needs to be aware of these dimensions and of their role in it. I think we also need to establish a high standard of transparency in AI, not only for open science purposes, but also for the purpose of an ecosystem of Public Interest AI and research on AI.

Sascha Schönig: Which role will AI play for science in the future? Are there any relevant use cases you can think of?

Theresa Züger: Most likely new tools for researchers will appear, they will improve and be adjusted to the needs of a specific research community. As I said, all-purpose solutions like ChatGPT are so bad at most things, where particular context specific details need to be considered, so let's put this one away. I don't see any time soon that AI will be able to replace the work that researchers are doing because so much of our work is about bringing contexts together, analyzing problems in-depth, not only superficially, considering important field-specific and historical developments. There will be tools that can support this kind of work. And I hope that they will come from the research community and adhere to our standards.

Because besides all the playing around, I think for academia, it's really important to stick to our values of producing knowledge, to make transparent, how we reached it and to have validity that we can prove. There is reproducibility in our knowledge, which is very important. I think we need to stick to these values and then tools can certainly help if they perform to assist us in achieving those goals. However, as I already pointed out, I see a much bigger potential in developing tools for specific areas of research where AI can help in addressing a niche problem, where researchers are using and training it themselves. This is where the real potential of such a technology lies, in my opinion, not in general purpose for tasks such as writing a student essay.

Sascha Schönig: I'm coming to an end so have I missed anything or you'd like to talk anything regarding the topic of AI in science or research?

Theresa Züger: One question that people always ask is, should students be allowed to use this and what should universities do? And I'm not saying there is a perfect answer to that, I like that many universities are not running away from that challenge, but rather trying to face it very openly and admit that we have to teach students that if they use it, how to use it smartly and be transparent about it, also learn to be aware of how the technology behind it works and how it's limited. Also, I think it challenges universities again to think about what their goals are as institutions. What do they want their students to be capable of when they get their degrees? And do we need to change the ways we do tests or how we make sure that the learning goals that we have are achieved?

INTERVIEW

And I think that's actually a very interesting task for universities to be challenged to do, to ask themselves from time and time again what they want to achieve and what kind of education and knowledge and training they want students to have.

Some universities are going back to oral exams, some others will figure out other types of tests or other types of education, interaction and how they want to make sure that people don't only replicate knowledge but really learn something. Make sure they are capable of using this knowledge, because that is the thing that universities are supposed to do. They're educating and they're teaching capabilities to people. And I also believe that most students, even though they might have a lazy day once in a while, are very much aware that they will only succeed and benefit from their education if they actually learn and enhance their capabilities. So the limitation to use these tools is hopefully also an intrinsic one, because you're not learning the same thing, and you're learning with limitations. you It definitely means we have to vary the way we learn.

Sascha Schönig: Theresa, thank you for your time and those insights.