Andreas Speer

AI and humanities

Digital humanities

There is no doubt that AI has become an integral part of the humanities. This comes as no surprise, as computer-assisted work has been firmly established in most humanities and cultural studies for some time now. This applies to object-based disciplines, such as the various branches of archaeology, as well as to text-based disciplines, such as linguistics, translation studies and textual criticism, but also to collections and archives of all kinds. Whilst the initial focus was on the digital documentation of sources and research findings and making them accessible as a consequence of the media revolution, digitality in all its facets now shapes the overall design and methodology of a research project from the very outset. Digitality is therefore an integral part of research design and makes certain research questions possible in the first place, both quantitatively and qualitatively – for example, through the co-presence of objects in archaeology and textual criticism and editing on one's own digital desk, regardless of where they come from or where they are currently located. And large corpora of linguistic and textual data form – not only in quantitative terms – the basis for new research questions and methods. Thus, for the humanities, databases are often the new libraries, even if – and this is a major challenge – the issue of sustainability remains unresolved. For if I re-edit a text written by hand more than 800 years ago, which is available in a 500-year-old printed Renaissance edition, on the basis of all available textual witnesses, then this edition, in turn, should be available for at least 500 years, just like its predecessor!

Acessability and sustainability

This highlights a distinctive feature of the humanities, and thus also of the digital humanities: the time horizon for research data is not so much determined by the immediate resolution of a problem, but rather by the long-term nature of cultural memory, which does not lose its value as time goes on. Quite the opposite. This is not merely a matter of storage, but of the accessibility and functionality of research platforms for users. Even AI will not solve this problem, even if we 'feed' it all the available texts – which is currently being done on a large scale, often at the cost of destroying the digitised objects themselves. But can AI really preserve our cultural heritage in the long term, and if so, how does it do so?

Prof. Dr. Dr. h.c. Andreas Speer
Universität zu Köln
Thomas-Institut
andreas.speer@uni-koeln.de

DOI-Nr.: 10.26125/19ad-8r58

This is evident not least in one of the key issues: search. More and more search portals are being infiltrated by AI, without it being clear where and to what extent the data and information are AI-generated. Information, however, must be verifiable. This requires not only the sound judgement of those receiving the information, but also transparency on the part of the body providing it. Particularly where the focus is not on technical or design-oriented solutions but on context-dependent understanding, constant reflection is required on (essentially unavoidable) biases, on preconceptions, on framework conditions, on interests and on the perspective chosen in the specific case.

For science – and this applies all the more to the humanities – cannot be understood as a supposedly objective, factual core that needs to be identified and defined, surrounded by a loose ring of social and cultural contexts that are disregarded when defining science. Rather, it is the interplay of contextual conditions – the available (or lacking) resources, the institutional and political sphere, alliances and the public sphere, indeed the mobilisation of the world – that constitutes the very object of science to be defined; this object is the result of a practice of interaction within a network of links and nodes that we can observe and study across a wide variety of epistemic fields.

Knowing why

This is evident in the current debates surrounding the use and, above all, the further development of AI. The calls for a moratorium on the further development of generative pre-trained transformer models – the GPTs or LLMs – or even for a complete halt to the development of this technology highlight a number of fundamental problems which, however, apply to any technology – albeit on a different scale: lack of transparency, loss of control, amplification effects caused by interconnection and automation, etc. On the one hand, we appreciate not having to rely constantly on instruction manuals in our daily lives, and are satisfied that, for example, my kitchen appliance works without us knowing exactly how. This illustrates the role technology plays in making our lives easier. Yet what may appear to be an extension of my scope for action actually diminishes my ability to cook for myself at all. I don't really know how the dish on my plate came about unless I've prepared it myself at some point and discovered in the process just how good it tastes when I deviate from the pre-programmed standard recipe.

Like any artefact, AI also focuses on very specific abilities that are intended to be enhanced by means of that artefact. In our case, it is not the flint of prehistoric man; rather, it is specific cognitive operations which, in turn, have a long history – without this implying (as is the case in some poor analogies) that

AI replicates the complexity of human thought processes, or indeed that AI thinks. Rather, the success of AI lies in its specific form of intelligence, which is based not on understanding meaning but on a complex form of pattern recognition. AI's strength therefore lies in its focus and specific power when dealing with a multitude of objects that are simultaneously present and processable, and in the speed with which this occurs. This characteristic of technology – its inherent ability to exceed its originally intended scope by a wide margin – can lead to a loss of control over the tool – particularly when we cannot comprehend the design blueprints and workflows, when we do not know why something is the way it is, and why something happens. However, the more powerful a technology is, the greater its impact, and the more urgently we must ask for the reasons – the 'why' – and must not be satisfied with the 'that' and the 'how'.

Writing and translating

This now seems to apply equally to the domain that characterises the humanities like no other: writing. Writing is more than just the production of text. Writing and thinking are inextricably linked. The writing process is often a laborious attempt to connect thoughts in a way that enables me, first and foremost, to understand a problem, a question, or even the outcome of my own thought processes. In doing so, I assess for myself whether I have understood something. Delegating the function of writing to AI restricts this capacity for understanding to a degree that is, in turn, scalable. The less I am able to determine the production of a text – or at least to follow it – the less I understand what that text is about. I essentially have no idea what criteria an AI uses to generate a text, or what sources it draws upon. Hallucinating is merely the most obvious problem, which is also linked to the bias of most chatbots to present a solution under any circumstances – even if it means providing a fabricated answer. And even in the field where AI is particularly powerful – translation – one is reliant on at least some knowledge of the languages in question in order to check the quality and appropriateness of the suggestions offered and to correct them where necessary. For token-statistical probability is blind to meaning and sometimes misses the context, which can only be grasped through an understanding of the meaning. Nevertheless, the language programmes and translation tools are impressive. They demonstrate the unique power of language corpora and the speed of search. Our old dictionary, whose columns we laboriously work our way through, doesn't stand a chance.

Shifting skills

The example of translation illustrates how AI is shifting the nature of skills required, even in the humanities. To stick with this example, translation is shifting from a basic task involving the use of a dictionary and grammar to the evaluation of suggested translations – which we should not simply accept, but rather evaluate and assess. The same applies to search, to suggested summaries, and to proposals for structuring a text. So what skills do we need to teach?



A recent article in the British business newspaper „The Economist" reported that university graduates with degrees in philosophy are among the few new entrants to the US labour market whose prospects have improved since 2022, when ChatGPT was launched. Furthermore – according to „The Economist" – major AI developers are increasingly seeking out philosophers, on the one hand to improve the 'thinking' processes of their language models. On the other hand, there is a need for suitably trained staff, for example to address ethical issues relating to AI. Perhaps philosophers (though not exclusively them) come up with ideas that do not arise from a recombination of existing data, but rather conceive something truly new. Above all, however, it is a matter of sound judgement and the responsible use of AI. For the responsibility lies not with the technology, but ultimately with those who invent, develop, deploy and use it.

References

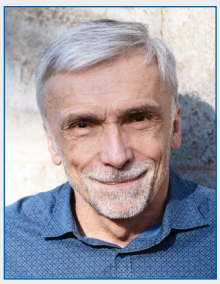
[1] N. K. Hayles, Bacteria to AI. Human Futures with our Nonhuman Symbionts, The University of Chicago Press 2025.

[2] S. Krämer, How should the generativ power of Large Language Models (LLMS) be interpreted? Do chatbots understand linguistic meaning?, in: *Philosophy & Digitality* **Vol. 2 No. 1**, 48-67 (https://doi.org/10.18716/pd.v2i1.11660).

[3] B. Latour, Pandora's Hope. An Essay on the Reality of Science Studies, Harvard University Press 1999 (dt.: Die Hoffnung der Pandora. Untersuchungen zur Wirklichkeit der Wissenschaft, Frankfurt a. M. 2002).

[4] A. Speer, Accessability: A Digital Perspective, in: *Philosophy & Digitality* **Vol. 1 No. 1**, 79-84 (https://doi.org/10.18716/pd.v1i1.2431).

Prof. Dr. Dr. h.c. Andreas Speer



Andreas Speer is professor for philosophy at the University of Cologne and director of the Thomas-Institut, an international research center for the history of philosophy and the sciences for a millennium we call the Middle Ages. He directs major research projects and conducts the biannual *Kölner Mediaevistentagungen*. He is speaker of the Cologne Center for eHumanities (CCeH) and Task Area Lead ‚Editions' of the NFDI-Consortium Text+. Since 2000 he is member of the senate of the DFG.