Artificial Intelligence as a cultural technique

Sybille Krämer in conversation with Jens Schröter, March 5, 2023

Schröter: You published a volume called "Mind, Brain, Artificial Intelligence" back in 1994. How has your view of so-called 'artificial intelligence' changed since then?

Krämer: I was – and still remain – convinced of the culturally shaped exteriority of the human mind: Having a brain is a necessary, but by no means the sufficient condition of our cognition. To think it is not a purely mental process in the head but is characterized by three other aspects: (I) the use of language and tools, (II) the social interaction with others, and (III) our corporeality and metabolism-based embeddedness in the ecosystem of our planet. This is the horizon in the 80s/90s when Artificial Intelligence (AI) aroused both fascination and criticism in me.

The fascination was based on the fact that rule-based symbol processing in the form of 'symbolic machines', which was practiced as a human intellectual technique long before the invention of the computer – for example in written calculation or logical deduction – always characterized a subarea of human problem-solving. To see how far machines with this paradigm of symbol processing can be developed – in the 80s these were the Expert Systems as a spearhead – does not reveal how human-like these machines work, but vice versa how machine-like humans have organized and still organize some domains of their cognition. So the remarkable fact for me about the then prevailing form of AI was not at all that computers can model the brain (according to the formula brain and mind like hardware and software) but that they adapt or simulate a cultural-technical practice, namely the handling of written symbols. It is not by chance that Alan Turing (1950) explicitly makes the human calculator, which enters, rearranges, and deletes symbols on checkered paper, the model of his mathematical-technical concept of the Turing machine. The difference is that the checkered paper has now become an endless tape.

On the other hand, my criticism was directed towards the myth of 'disembodied intelligence', associated with the symbol-processing approach of AI, as soon as this is generalized as a model of human thinking and our being-in-the-world. This was one of the critical arguments of Hubert Dreyfus (1972), thereby going back to Heidegger. By virtue of our bodily situatedness, we have a primordial relation to the world that is independent of explicit symbol processing, a pre-symbolic intuitive understanding that implicitly structures our practices. Then, something came to the fore that marked the limits of the symbol-processing paradigm.

This was roughly the tableau of my initial involvement with AI at the end of the last century.

However, with the mass data made possible by the Internet, social platforms, and ubiquitous computing – used to train artificial neural networks, especially in Deep Learning – the role of AI in society has fundamentally changed. Here are some symptoms of this change:

I. Artificial Intelligence has arrived in everyday life (search engines, face recognition, spam filters, navigation, chatbots, etc.) – in other words, it is not only used as a selective expert system. In everyday applications, it mostly remains hidden from users, often – though not always – operating below the threshold of perception. This is changing with Large Language Model-based chatbots, which respond to colloquial prompts and thus advance to an everyday technique.
II. The increased generative potential is conspicuous since both images and texts can be generated today with natural language instructions, each of which is unique, not plagiarized.
III. Since learning systems are no longer explicitly instructed via programming, but are primarily trained by sample data and error feedback, the internal models formed in the process remain opaque: The area of non-knowledge in systems, that nevertheless function well, is growing.
IV. Early AI was seen as a tool to uncover functional processes of the human mind (= symbol processing) or brain (= connectionism). Now, learning algorithms teach us about the discriminations implicit in our social practices that condense into training data. By practically executing biases represented in training data, algorithms at the same time bring them into the open.
V. Statistical language analysis and language generation have superseded attempts to model semantics, meaning, and comprehension. The ap-

proaches of the Large Language Models, especially the 'family' of ChatGPTs, show: What the machine generates is not based on understanding, but on the statistical combination of elementary tokens (small groups of letters below the level of meaning) according to the most probable linkages. Thus, the astonishment, in how many respects ChatGPTs can produce plausible texts, corresponds to the insight that precisely no intelligence is required for this. What is necessary, however, is combinatorial access to billions of texts – which is not feasible for humans – in order to create products whose reference to reality is fictional – i.e. without any claim to truth. Does quantity – the unsurpassable large training data volumes – turn into quality here? Or has the demarcation line between quantity and quality become questionable in general?

Schröter: How would you classify the development of so-called 'artificial intelligence' in the history of formalization that you have studied in detail? Today's dominant machine learning methods belong to a rather statistical paradigm – does this belong to the history of formalization or rather not?

Krämer: Formalization does not mean calculating with numbers, but manipulating graphic signs according to given rules. The philosopher Leibniz first articulated this distinction (Krämer 2016). In written reckoning, the eye, hand, and brain work together and create a 'machine room of intelligence' that consists of formal pattern manipulation and is independent of using a real physical machine. The signs can represent numbers, but they do not have to. The procedure itself is an interpretation-independent operation of forming and transforming strings of signs. In memory of handwritten calculating: If a table with one and one, one minus one, one times one, one divided by one is available, then elementary arithmetics can be carried out with paper and pencil, without having to know at all that numbers are processed. This, at least, is the sense of formality that emerged with the development of mathematical and logical calculi in the modern era. Of course, formalization has no end in itself: If a consistent object domain is discovered as a reference domain of a calculus, domain-specific problems can be solved formally and new insights can be gained.

This being said, any operation with numbers, regardless of how the calculation is performed and whether probability and statistics play a role in it, is necessarily formal. How formality and statistics are related is exposed when the sentence is correctly understood that in 2021 each woman in Germany had 1.58 children.

But we had to add another dimension with regard to the relationship between machine learning/statistics and formalization. It is the transition from problem-solving to predictive algorithms, which is crucial for contemporary digitization. Problem-solving algorithms determine a result in a stereotypical mechanical way: By applying the rule of calculation correctly, the result will be correct too. You can 'trust' the algorithm. Predictive algorithms, on the other hand, refer to the future and predict the probability that a possible event will perhaps occur. Already in the case of problem-solving algorithms, the 'knowing-that' splits from the 'knowing-how' in the application: The knowing how to do something becomes transparent, teachable, and learnable; the knowing why it works remains hidden and is at best transparent to mathematicians, but not to the calculators.

In contrast, in predictive algorithms, the machine acquires a knowing-how in the form of an internal model, i.e. the functional competence to make an input correspond to an output. The 'knowledge' implicit in this internal model usually cannot be inferred from the output and remains opaque; apart from that, these internal models change with every use and in innumerable permutations. Moreover, with predictive algorithms, the social and political importance of the presupposed labelling grows, i.e. the mostly human selection and marking of training data as well as the social scaling of thresholds in the internal model building.

We see: Every algorithmization implements and embodies a specific relationship of knowledge and non-knowledge, of transparency and opacity; but in predictive algorithms, the domains of non-knowledge and uncertainty radically increase.

In view of this situation, doesn't the idea of 'Explainable AI' also create an illusion? Do we perhaps have to radically change our attitude and perspective with regard to the relation between knowing and not knowing? Is it not rather a matter of reopening the fundamental questions of knowledge/non-knowledge, of acting under uncertainty, and all this in the opposite direction too: A medical doctor interpreting an X-ray is much more likely to act under the sword of Damocles of uncertainty than a system trained to make these diagnoses with thousands of analyzed X-rays. Are common terms like 'knowledge society' emphasizing enough that every new knowledge creates new not-knowing, that we cannot always eliminate uncertainty but have to learn how to deal with it? And that human action cannot escape this ambivalence?

Schröter: How would you relate to the development of so-called 'artificial intelligence' in contrast to the somewhat fuzzy discourse of 'digitalization'? How would you relate to the assumption, that at least neural networks are rather analog technologies, again because of the finely graded weighting of the activity of artificial neurons, and because of their parallelism (cf. Sudmann 2018)?

Krämer: The digital exists – this may come as a surprise – before and independently of the computer. By digitization, I mean a process in which a continuum is broken down into basic elements and discretized so that they can be coded and combined with each other in a more or less arbitrary way. A prototype for digitization is the alphabet. Although the flow of oral speech knows breaks, they do not correspond at all to the blank spaces between words and sentences in alphabetic writing. With the finite repertoire of alphabetic characters, an unlimited number of combinations can be produced in the two-dimensionality of a surface. This non-linear 'nature' of writing is revealed for example by the phenomenon of the crossword puzzle which exists only as a two-dimensional, graphic medium illustrating the novel configurations that spatial writings open up in comparison to temporal speech. Moreover, alphabetically ordered lists sort large amounts of information, think of the traditional telephone directories, which allow casual access to amounts of data that cannot be surveyed by humans. A 'database principle avant la lettre' developed in social practice is already being applied: the abandonment of narration in favor of formal sorting and addressing of pieces of information that are independent of each other. This database principle gave rise to the academic flagship projects of print-oriented modernity in the form of dictionaries, encyclopedias, and lexicons.

Let us summarize. Two things are important with regard to my concept of digitization:

(1) There is an embryonic digitality already connected to alphanumeric literacy. This does not only apply to the European alphabetization: The hexameters of the Chinese Book of Changes "I Ching", for example, are written with dual code, which can be translated without constraint into machine-processable Unicode. The digital is to be understood independently of computer use.
(2) The relationship between analog and digital is relative in so far as we understand it in terms of the continuous-into-discrete transformation. The transition from fluid speech to discrete writing is a transition from an analog

> to a digital medium. But if the transformation from a printed text to a machine-readable and -analyzable document encoded in TEI, is considered, then the printed typeface is in the role of the analog and only the encoding instantiates the process of digitization.

In a significant way, the connection between digitality and Artificial Intelligence is clarified by their latest development: The already mentioned contemporary chatbots in the context of Large Language Models (GPT-4, Bard etc.) operate on the basis of small, meaningless groups of letters, the 'tokens'. Here, too, we are dealing with the decomposition of something continuous into smaller meaningless units. Hardly anything can better illustrate how 'deeply' the techniques of Artificial Intelligence are allied with the digital, understood as a process of dynamic discretization.

It should be recalled that linguistics characterizes human language by its 'double articulation'. From a limited repertoire of meaningless elements such as phonemes or letters, an unlimited number of meaningful words and sentences can be formed. The question arises if a digital principle is already nested in spoken language – at least implicitly. However, there are good reasons to assume that the phoneme is the result and product of the grapheme, the smallest written unit. In fact, only the emergence of phonetic writing has split and divided communication in its totality of prosody, mimic, gesture, deixis, and verbality and crystallized the phonetic dimension as an independent communicative strand and condensed it to an object like perceivable 'language'. If this is true, it would be the writing that puts the grid of digitizing over human language.

And one last remark: If your question aims at a possible return of the analog by artificial neural networks, I am skeptical about any neuromorphic diction and rhetoric. Bird flight also inspired human flight experiments, without airplanes imitating the natural model. Is it not the same in relation to natural and artificial neural networks? Everything that matters in contemporary Artificial Intelligence, is mostly not programmed but trained by huge databases, and what can explain its technical power is something that finds no role model in nature. The procedure of error feedback, for example, which has an analog in the social practice of teaching when corrected dictations are returned, finds no parallel in neurophysiology. Or with regard to the architecture of the hidden layers – a central component of the Deep Learning process: If each layer analyzes selected aspects of the input with different weighting, or if these computational processes take place in the layers one after the other – all this also has no analog in our brain. Not to mention, by the way, the energy efficiency that is so typical for our brain.

Schröter: What role do you think methods of so-called 'artificial intelligence' could play in the field of digital humanities? How could machine learning be used in the cultural sciences and humanities, and even in philosophy?

Krämer: In this context, I'd like to talk about the 'sting of the digital'. What I would like to express here is that the debate about the Digital Humanities and their acceptance by the traditional humanities can provide impulses for a self-correction of the humanities' self-image. This self-correction refers to the absolutization of hermeneutics and interpretation as the royal road and definiens of the humanities (Krämer 2023). Furthermore, using 'sting' as a metaphor refers to criticizing the belief that the humanities have nothing to do with empiricism or with material and quantifiable things and processes. Incidentally, both of these biases have already been subject to erosion in the late last century, even independently of the emergence of Digital Humanities.

The humanities' disciplines encompass not only the traditional fields from history to linguistics, literature, music, and art studies, but also archaeology, ethnology, and even regional and cultural studies. They have always worked with materials, that is, with things, documents, and artifacts of all kinds, which are to be collected, dated, classified, annotated, compared, archived, and so on. In this ecosystem of scholarly work in the humanities, empirical questions – and thus numbers and counting – always had a certain status. But the traditional humanities with their hypostasizing of interpretation as key methodology, have long remained blind to the materiality of their research objects and consequently to the importance of numbers and countability in many subfields of their research.

Nevertheless, it is precisely here that research questions open up that can be meaningfully addressed by the Digital Humanities under the conditions of contemporary digitization. This is always the case when large data corpora, which relate to lifeworld and/or cultural-historical contexts and can no longer be surveyed, let alone examined, by human eyes and hands, can now be analyzed with data-driven, computer-based methods. However, this is only possible through the subtle, difficult, never-ending interaction between researchers and computer-generated, data-driven procedures. It goes without saying that interpretation on the part of human actors is constantly involved: no number – and no data – interprets itself.

In prosaic terms, the question of sense and nonsense of the Digital Humanities could be transformed into the question of what role empirical questions play in the respective discipline. Against this background, it is not surprising that datafication and digitization first took hold of the natural sciences and, in the 20th century, also of economics and the social sciences, before it has now arrived in the 21st century humanities. Perhaps the discussion about the legitimacy of Digital Humanities serves as a proxy function for the less exciting question of when and how the empirical can or should gain a birthright in the humanities.

We must not make the mistake of reestablishing C.P. Snow's two-culture difference (Snow 1959), which is unacceptable today, within the humanities. Even the traditional humanities have always been dependent on dealing with numbers and data, think of concordances that have existed since the 13th century, catalogs of works or historical dating, etc., just as, conversely, the Digital Humanities always have to interpret their results in the light of their research questions. There is no such thing as interpretation-free empirics.

In the opposite direction, however, I also find problematic contemporary attempts to identify and ennoble computational procedures themselves as hermeneutic procedures, as Dobson did in 2019, for example, in order to provide the Digital Humanities with legitimacy in the Humanities. As already emphasized, I am more inclined to weaken the hermeneutic paradigm as a unique selling point of the Humanities by recognizing that their academic practices include a plethora of activities in the preparation of their research objects that precede and prepare the ground for interpretation in the first place.

However, there is an interesting and revealing addition to this statement. Computers are forensic machines (Kirschenbaum 2012), like microscopes and telescopes directed toward the data universe to find patterns that mostly escape human attention. Of course, the optical analogy is limp insofar as it ignores the generative aspect of processing and synthesizing music, images, and text. However, what is at stake in explaining the forensic function is the dimensions of the culturally unconscious. What people miss in their practices, a machine can register.

This can be explained by the computer-philological example of author attribution. If styles of individual authors become identifiable by means of a ranked list of the 'incidental' functional words used – how often are words like 'and', 'nevertheless', 'however', etc. being used? – then the machine is able to identify an author by attributes of his or her use of language that is not at all part of the stylistic devices intentionally employed, but rather is subverted in writing and occurs unconsciously in the performance of written articulation. It is not about something that is hidden behind what is written, but that is given in what is written down. It is implicit in the surface of the text and can therefore be taken from it.

What emerges here within the dimension of author attribution is generalizable: Despite the use of terms such as 'Deep Learning', information processing technologies – also in the form of Artificial Intelligence algorithms – are a surface technology for the identification, analysis, and production of patterns. What is true for numbers and data is also true for patterns: Whether patterns have meaning, sense, and relevance, be it for life or for a research question is up to humans to decide, applying the pattern discovery capacity of the machine for their specific purposes.

It has hardly been registered so far that 'close' and 'distant reading' converge in this question. The cultural scientist Carlo Ginzburg (1983) – as a micro-historian, he was an advocate of close reading – saw a 'circumstantial' or 'indication paradigm' emerging as a methodological dispositive of the humanities in the transition from the 18th to the 19th century. The inventor of the detective Sherlock Holmes, the author Arthur Conan Doyle, the art historian Giovanni Morelli, and the psychoanalyst Sigmund Freud developed their insights by studying unnoticed details at crime scenes, in faked paintings, and in traumatized souls. In this way, Ginzburg was able to show why Doyle's detective novel became the most successful crime novel series: because readers are involved in the process of finding clues. The propagandist of distant reading, Franco Moretti (2013), in turn, by comparing all detective novels in Doyle's epoch (a fact Ginzburg could not have had an overview of), comes to a very similar conclusion, namely that of the exceptional position Doyle's "Sherlock Holmes" novels had.

The micro perspective of close reading and the macro perspective of distant reading are not opposing perspectives but can complement each other. Furthermore, something else becomes clear here: Statistical methods are often reproached by the humanities because they only represent the average and are therefore an instrument for the enforcement of mediocrity and the renunciation of creativity. However, statistically operating computational methods do not only calculate average and mean values, but by virtue of this computational capacity they can also uncover the knitting pattern of the individual from a most unusual perspective, just as forensics can uncover a singular course of events or author attribution can uncover author identities. However, this always works only probabilistically, i.e., by a probability statement. In short: Statistics is not the enemy of casuistry and of the individual case, but – used sensibly – can be precisely its aid.

Schröter: Can so-called 'artificial intelligence' be described as a 'cultural technique'? Or does it rather presuppose certain cultural techniques?

Krämer: Every technology is socially constituted and thus a cultural phenomenon. And yet a distinction must be made between 'technology' and 'cultural techniques'. In the context of the in 1999 started Helmholtz Center for Cultural Techniques in Berlin – I was a member of the eight-member founding group – the term 'cultural technique' aimed to orient research in the humanities more strongly towards the materiality, mediality, and technicality of their research objects. In this Helmholtz group, cultural techniques were regarded as routinized everyday procedures for dealing with symbolic and technical artifacts that are sedimented in everyday practices, the mastery of which provides a basis for social participation, but also for social differentiation. Cultural techniques are crucial resources of scientific and artistic practices and also underlie higher-level cognitions.

We are familiar with the fact that writing, reading, and calculating are cultural techniques of the era of printing. From this point of view, it is obvious that digital literacy implies a decisive development of those cultural techniques that have been typical for alphanumeric literacy in the 'Gutenberg Galaxy'... The elementary handling of keyboards, smartphone use, the ability to communicate by email, and, above all, to search for information on the Internet are decisive aspects of contemporary digital cultural techniques, without which participation in social life is hardly conceivable. At the same time, these are practices at whose mastery or non-mastery fault the lines of contemporary society emerge, both socially, but also generationally. But does this also include the processes of Artificial Intelligence?

For the era of Expert Systems – i.e., in 'woodcut' terms: the AI of the last century – I would have answered this firmly in the negative. But precisely because contemporary Artificial Intelligence has seeped into our everyday behavior in many different forms, the situation has changed. Without streaming, navigating, searching the net, online banking, spam filters, etc., contemporary participation in everyday life seems almost impossible to realize – although in principle this remains possible, just as illiterate people can lead a special existence in literal cultures. This dependence on the cultural techniques of Artificial

Intelligence also applies to complex mental work: Without computer-generated visualization, medical diagnoses and operations are hardly feasible anymore, stock market trading thrives on real-time analyses, driving assistants in cars have become standard, and fitness watches control training and mobility. A significant step in everyday usability of AI is the software trained with large data corpora, allowing users to instruct image and text generation with natural language – and its colloquial character is important.

However – and this also seems to be a novelty in the degree of the associated dangers – Artificial Intelligence procedures often run as background processes that are hardly registerable for users, let alone recognizable and accessible. In a harmless dimension, when taking photos with a smartphone or in the use of auto-correction functions, but more problematically in the creation of personal data profiles as 'waste products' of Internet navigation.

Artificial Intelligence nowadays is implemented into the use of apps, objects, and procedures. The cultural technique consists in being able to deal with virtual objects in a functionally and factually appropriate way without having to understand how this use of data can be exploited in a functionally and factually non-intended, but commercialized way. What I have characterized as the dispositive of technology use – i.e., being able to control and use without having to understand – acquires an ethical-political signature here. Can we conclude from this that the cultural technique of Artificial Intelligence also consists in learning how to preserve data sovereignty? Or is this idea of sovereignty, rooted in the European Enlightenment with its maxim of 'thinking for oneself', an illusion – and perhaps was from the very beginning? For it is precisely the suitability of these everyday applications which become smarter with each use, that is in turn restricted, if not hindered, by mechanisms of data protection: Who isn't annoyed by the popping up of the cookie consent form, which degrades data sovereignty to check-marking? How much more helpful could digitization be in Germany if patient data or even the data available in administrations were merged? A dilemma is emerging between smart everyday usability and responsible handling of Artificial Intelligence's 'background cultural technology'. 'Dilemma' is understood here as a conflict situation and a predicament that cannot simply be transformed into a positive solution.

Schröter: Would you see the use of machine learning in different sciences as a kind of upheaval – or rather as a continuation of the increasing role of computers in the sciences (e.g., in the form of computer simulation)?

Krämer: Wherever the dynamics of media innovations are concerned, they are always to be understood in the tension between continuity and breakup, between tradition and disruption.

To give a distant example: The absence of book religion in ancient Greece allowed written texts to advance into a non-canonical discursive space debating the pros and cons of truth claims. What was previously known only from the oral practices of court proceedings in Greece, was now transposed into a written medium. Thus a type of text emerged, often in dialogue form as in Plato, which insisted on arguing about truth – and this became a relevant starting point for the Western type of philosophizing. This change is often called the transition from orality to literality, a highly problematic thesis, in whose garb mostly the Eurocentric assumption of the superiority of alphabetic writing was transported. Of course, orality is not replaced and made obsolete by literacy. Rather, writing opens up a symbolic space in which new ways of using and dealing with language become possible. And the oral also takes on new signatures, for example in the genre of the scientific lecture.

But back to the digital: Undoubtedly, the computer is currently becoming a universal tool in the sciences, from simple word processing to computer simulation. I use the word 'computer' here as a chiffre for the ecosystem of scientific information processing based on ubiquitous datafication. To stay with computer simulation, it is not simply that computer simulation now joins experiment and theory as a third research pillar in the sciences. Rather, this simulation opens up a new kind of mediation between analytical theory and empirical experiment: Experimenting with theories becomes possible (Gramelsberger 2008) and gives rise to a 'theory laboratory'. Computer simulation opens up a space in which traditional instruments of knowledge such as theories and experiments gain a new profile, combined with new options for knowledge.

Under the conditions of extensive datafication on the one hand and 'learning algorithms' on the other, this new profile is that computers can work with mass data in ways unattainable by human power. The forensic capability of computers, familiar with criminalistic use, can now be extended to many areas of scientific research, where it can be used to uncover patterns that are beyond human perception.

If the computer acts like a microscope and telescope on datafied worlds in data-driven research methods, then data corpora reveal and uncover what remains invisible to limited human perception. These computer-processable traces are mostly statistical, hence numerical constellations. And since neither traces nor data and certainly not numbers are self-interpreting, it is clear that only the research motivation, creativity, and synthesis of human interpreters can produce meaning and content from these traces, data, and numbers. Humans combine computer-generated results with theses, theories, and narratives and thus turn data processing into knowledge production.

Therefore, the question of the relationship between continuity and upheaval, between continuation and innovation in the scientific use of computers must be answered with a 'both/and' – as is usually the case with disjunctive questions.

The continuity of the development is unmistakable: It is well known that machine learning and the imitation of the human nervous system played a role already in 1956 at the conference at Dartmouth College, where McCarthy introduced the term 'artificial intelligence'. Turing had already raised these questions in the 1940s. In 1957, Frank Rosenblatt conceived the first artificial neural network with the Perzeptron; in 1966, Joseph Weizenbaum created the first chatbot with Eliza – and shook up the humanities scholars at the latest as a result of the illusion evoked by users of Eliza that an empathetic human was speaking here. Over the years, many other stations were added: Expert Systems in medicine, oral speech synthesis, winning chess, Go and quiz programs, chatbots such as Siri and Alexa, and finally, the image and text-generating artificial neural networks based on Deep Learning methods, training, and testing: Artificial Intelligence – regardless of its many slumps and crashes in the public consciousness and the seasonal metaphors like 'winter of artificial intelligence' that are readily used for this purpose – forms Ariadne's thread in the history of technology and science of the last decades.

Nevertheless, there is also an innovative, disruptive dynamic – and its symptom is the cultural-technical embedding of Artificial Intelligence in everyday practices. This cannot be monocausal traced back, for example to the use of Deep Learning processes from around 2012, but includes at least two other indispensable components: the datafication, doubling our world into the shadow image of a computer-processable data universe, and the extremely increased computing power of the hardware. The Deep Learning procedures become better and better with each increase in the amount of data – which was not true of machine learning in the early days of Artificial Intelligence – and increased amounts of data, in turn, require increased computational power, and so on. From the swirling dance of these three conditions with each other, has now entered the family of Large Language Models to the public; this has already been interpreted as the 'iPhone moment' of Artificial Intelligence. It is also significant for interpreting Artificial Intelligence now becoming a cultural technique, that it was OpenAI enabling the download of ChatGPT for all interested people (100 million users after only two months). All big players in this field will go to the market with their own versions, and Microsoft already announced to incorporate Large Language Models into its Outlook and Office programs. Search engines – but they were that before, ergo: continuation and break!

Schröter: How can the already so-called 'artificial intelligence' be placed in the history of the 'exteriority of the mind', which you have been investigating for quite some time?

Krämer: We are familiar with understanding humans as meaning-giving and symbol-oriented living beings who constantly interpret their world. Who would and could contradict this? But does looking for meaning and interpretation take it all? Civilizations develop by increasing the areas structured in a way that is independent of interpretation, reflection, and understanding. This is true not only for formal operations in the context of intellectual work but also for ritualized everyday practices. We celebrate Christmas even without a Christian message, drive cars without an understanding of technology, cook without an awareness of chemical interactions, and successfully apply computational algorithms. Alfred Whitehead (1911) remarked laconically at the beginning of the 20th century that the level of development of a civilization is shown by how many of its important operations can be performed without thinking about them.

Let us note: The dispositive of the use of technology consists in being able to apply and control without having to understand. And exactly this technical dispositive is transmittable to subareas of mental work too.

In addition, there is the collective character of the mind: Humans do not simply have natural intelligence but participate in different degrees in the socially shaped and distributed mind, acquired, passed on, and handed down in the collective. Our cognitive capacity can only be reconstructed as social epistemology. It already starts with an almost trivial fact: 85 percent of what we know, we cannot verify and justify on our own, but we acquire this knowledge through words, writings, and images from others. And trust is that very bond, the 'glue' that turns received information into knowledge for us. Here, with the knowledge machines of AI, an important moral problem emerges: How far can we trust the apparatus and the algorithms? Not at all in the case of the ChatGPTs, which generate their plausible-sounding texts as purely fictional products without any reference to reality, without any internal truth check (work is being done to change this). These machines have no mind, and no understanding, but calculate the probabilities of small tokens and word patterns.

Back to the question of the extended human mind: Without the exteriority of auxiliary means, starting with spoken language, including the manifold forms of visual signs, up to ornaments, pictures, graphs, diagrams and maps, scientific cultures and other functional areas in society would be unthinkable. To paraphrase Ludwig Wittgenstein: Why do we say that our thinking is located in the head and why do we not say that the speaking mouth or the writing hand is thinking too? We do not think on paper, but with paper.

In the context of the human mind's evolution in the interplay of eye, hand and brain, the cultural technique of flattening plays a central role. Here, 'flattening' is not meant in a pejorative way, but rather in a sense that inscribed and illustrated surfaces embody an irreplaceable, often creative, potential as a workspace for designing, as a thought laboratory, or as a workshop for composition and combinatorics. Just as we use geographic maps to orient and move in unfamiliar terrain, the diagrams and graphs of science provide a cartographic impulse for orientating and operating in conceptual spaces of knowledge: invisible entities, and non-spatial abstractions become representable and processable in two-dimensional spatiality. Our conception of time is also rooted in this potential for spatialization; we need only to think of the historian's timeline or the measuring of time by clocks. The inscribed or illustrated surface as a medium in between temporal one-dimensionality and spatial three-dimensionality is a translation manual from time into space and vice versa. To avoid misunderstanding: There are no flat corporal objects empirically, yet we treat inscribed and illustrated surfaces as if they are two-dimensional. Given the diagrammatic practices of knowledge, we realize how strongly the computer and the digital are linked to the exteriority of artificial flatness.

This is not only true for computer programs, which have to be written down before they can be used as machine instructions; it is also true for the model of the Turing machine, which works with a tape that can move back and forth, or is true for the multiplication of surfaces, which is typical for the architecture of the 'hidden layers' in Convolutional Neural Networks, and it is not least true for all the visualizations that are necessary to transfer computer-generated outputs into a form that can be understood by humans. And this applies basally already to encoding in TEI: Implicit reading conventions that we master as tacit knowledge by distinguishing and recognizing headings, footnotes, paragraphs, and proper names from one another in a text, must be made explicit line by line when encoding into a computer-processable script. The computer is a surface technology; therein lies its power and its limitations. As a microscope and telescope into the data universe it is unsurpassed – but also only within the data universe. What is not in this universe, does not exist for the computer.

Schröter: In 1998, you published the beautiful volume "Medien, Computer, Realität" (Media, Computers, Reality). The subtitle was "Concepts of Reality and New Media". What 'conceptions of reality' are associated with so-called 'artificial intelligence'?

Krämer: The idea that explaining our brain is to think along the lines of computerized operations, i.e., that phantasm (of the beginnings of Artificial Intelligence) to assume that the computer is the appropriate model for the human mind, is taken ad absurdum precisely because the latest chatbots are based on Large Language Models. The fascinating range of text genres produced by chatbots is – as we all know – free of all understanding on the part of the machine. The machine does what it does best after being fed huge corpora of Anglo-Saxon training data to calculate probabilities of letter tokens and word combinations.

The idea that technical apparatuses and processes displace and substitute people is problematic too. What AI actually demonstrates, is that we have to understand the relationship between humans and technology as co-performance – as a shared activity and interaction. Could we go so far as to think of human/machine interaction under the precinct of contemporary digitization, according to the model of alternating moves that are performed in a game?

Therefore, the talk of so-called 'self-learning programs' is distorting. Even when a computer defeated the four best poker players in the world, the winning program Libratus still had to be trained at night during the competition by its creators on the basis of game data. Rainer Mühlhoff (2019) elaborated on the socially distributed nature of Artificial Intelligence by pointing to the work armies of cheap click workers whose job it is to label the training data. In processes like CAPTCHA, where we have to read distorted strings or to name image objects to prove and identify ourselves as human, we fill the pool for training data of learning algorithms in involuntary pandering.

Schröter: It has become a standard argument to criticize so-called 'artificial intelligence' on the one hand because of the 'bias' of the data sets and on the other hand because of the lack of 'explainability'. In your opinion, are there other important criticisms of 'artificial intelligence'?

Krämer: First, the short answer. There are at least 3 points of view:

I. The resource problem: Artificial Intelligence in the mode of artificial neural networks or so-called 'self-adaptive algorithms' require immense data corpora. Our data universe is not infinite. This is not only an ecological problem of high power consumption. It is also about the fact that the algorithms' appetite for data multiplies the options in terms of data abuse and raises questions about data protection, copyright, etc.
II. The history of Artificial Intelligence – starting with its name – is also the history of the use of distorting terms such as the talk of 'self-learning systems'. The degree of self-sufficiency that this term evokes does not exist. All algorithms made efficient by training owe their potential to depend on interactions with humans, whether in labelling data, in graduating parameters, in deciding when output is considered 'efficient enough', etc. Or this talk of data as 'raw material', which also resonates in the phrase 'data mining': Data are artifacts, even if the data are based on measurements of the real. They are human-made: manufactured, not found. Here, an alarming proximity to the idea of 'nature as raw material' comes into play and thus, to a worldview oriented towards the exploitation of nature, the limits of which we are now – sometimes dramatically – confronted with.
III. Finally, it is important to mention the dominance not only of the English language, but of the Anglo-Saxon cultural asset and heritage that goes into the huge training bases of contemporary Large Language Models. The queries and instructions possible in the national languages, as well as what the system provides colloquially, are based on (machine) translation.

Let's keep in mind: Mistakes of today's AI are the technical advances of tomorrow! For example, the metamorphosis into a racist led to the removal of chatbot Tay (released by Microsoft 2016) from the network, and this metamorphosis became an instructive topic of debate; similarly, BlenderBot (released by Meta 2022) mutated into a supporter of conspiracy theories. Learning algorithms mirror the practices on the basis of which they learn, as if through a magnifying glass: It is up to us to learn how to use the computer as an instrument of self-recognition – and not only in the form of the fitness bracelet. We should address Artificial Intelligence less from the perspective of modeling and technical projection of mind and intelligence, but more as a virtual mirror of human communication. Elena Esposito (2022) has convincingly argued that not Artificial Intelligence, but artificial communication, is the operational basis of current computer use.

A further, even more complex answer to the question of criticism suggests itself to me: Is the gesture of critique itself, which founded academic modernity, perhaps reaching its limits at present, as Rita Felski (2015) suspects?

The gesture of 'critique' is deeply anchored in the humanities' self-image of scholarly work. Unfairly shortened to the formula: Saying 'no' is always possible, saying 'yes' is under suspicion of apology. But are we as humanistic scholars really 'by profession' in the position of a meta-position towards that which we criticize? With the consequence that we are entitled to actually judge and evaluate from the superior standpoint of a knowledge that has both an affinity for technology and at the same time looks ahead to the future? Wasn't it precisely a concern of the convinced hermeneut Hans-Georg Gadamer (1975 [1960]) that humanities scholars should be regarded not in the bird's-eye perspective as observers, but in the participants' perspective as players in the events of the world, entangled in prejudices? Perhaps this is the reason why I do not focus on the critique of AI, but want to shake up the prejudices in which the humanities are caught when they take a stand on digitalization and Artificial Intelligence. To enlighten about technology means first to understand technology to some extent and second, to free its use from myths.

What is critical, is not so much AI itself as a technical endeavor, because we need technology to solve the problems of this planet in a way that can be both accepted and welcomed by the people whose behavior needs to change. Rather, what is critical, is our use and abuse of technical potential and the myths and ideologies surrounding it.

In fact, critical humanists like to focus on the ideologizations and mythicizations, apocalyptic and apologetic interpretations of Artificial Intelligence – and then often pass this off as a critique of AI itself or misinterpret it as such. I, therefore, argue for a kind of 'sobriety' in the discussion of AI. It is still about an – albeit interactive – 'toolbox', whose fields of application are growing by the hour, not to say proliferating.

It is not the intelligence and rationality of machines that we have to fear, but the irrationality of people.

List of references

Dobson, James. E (2019): Critical Digital Humanities. The Search for a Methodology, Champaign: University of Illinois Press.

Dreyfus, Hubert (1972): What Computers Can't Do. The Limits of Artificial Intelligence, New York: Harper & Row.

Esposito, Elena (2022): Artificial Communication. How Algorithms Produce Social Intelligence, Cambridge, MA: The MIT Press.

Felski, Rita (2015): The Limits of Critique, Chicago and London: The University of Chicago Press.

Gadamer, Hans Georg (1975 [1960]): Wahrheit und Methode. Grundzüge einer philosophischen Hermeneutik, 4th ed., Tübingen: J.C.B. Mohr.

Ginzburg, Carlo (1983): Spurensicherungen. Über verborgene Geschichte, Kunst und soziales Gedächtnis, Berlin: Wagenbach.

Gramelsberger, Gabriele (2008): "Computersimulationen – Neue Instrumente der Wissensproduktion." In: Renate Mayntz/Friedhelm Neidhardt/Peter Weingart/Ulrich Wengenroth (eds.), Wissensproduktion und Wissenstransfer: Wissen im Spannungsfeld von Wissenschaft, Politik und Öffentlichkeit, Bielefeld: transcript, pp. 75–96.

Kirschenbaum, Matthew (2012): Mechanisms: New Media and the Forensic Imagination. Cambridge, MA: The MIT Press.

Krämer, Sybille (2016): Leibniz on Symbolism as a Cognitive Instrument. In: Philosophy of Emerging Media ed. Juliet Floyd and James E. Katz, Oxford: Oxford University Press, pp. 307–318.

Krämer, Sybille (2023): "Should We Really 'Hermeneutise' the Digital Humanities? A Plea for the Epistemic Productivity of a 'Cultural Technique of Flattening' in the Humanities." In: Journal of Cultural Analytics 7/4 (https://doi.org/10.22148/001c.55592).

Krämer, Sybille (ed.) (1994): Geist, Gehirn, Künstliche Intelligenz: Zeitgenössische Modelle des Denkens. Ringvorlesung an der Freien Universität Berlin, Berlin and New York: De Gruyter.

Krämer, Sybille (ed.) (1998): Medien, Computer, Realität. Wirklichkeitsvorstellungen und Neue Medien, Frankfurt a.M.: Suhrkamp.

Moretti, Franco (2013): Distant Reading, London: Verso.

Mühlhoff, Rainer (2019): "Menschengestützte Künstliche Intelligenz. Über die soziotechnischen Voraussetzungen von 'deep learning'." In: Zeitschrift für Medienwissenschaft 11/2, pp. 56–64.

Snow, Charles Percy (1959): Two Cultures. The Rede Lecture: Cambridge: Cambridge University Press (https://www.rbkc.gov.uk/pdf/Rede-lecture-2-cu ltures.pdf).

Sudmann, Andreas (2018): "Szenarien des Postdigitalen: Deep Learning als MedienRevolution." In: Christoph Engemann/Andreas Sudmann (eds.), Machine Learning – Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz, Bielefeld: transcript, pp. 55–74.

Turing, Alan (1950): "Computing Machinery and Intelligence." In: Mind 59/236, pp. 433–460.

Whitehead, Alfred N. (1911): An Introduction to Mathematics, London: Williams & Norgate.