Research with Subsymbolic AI
Preliminary Theses

Andreas Sudmann, Anna Echterhölter, Markus Ramsauer, Fabian Retkowski, Jens Schröter

The current developments within information technology not only challenge scientific disciplines to study new phenomena; they also potentially alter and enhance research methods, practices, and outcomes across the natural sciences, social sciences, and humanities. Researchers have to negotiate interdisciplinary conceptual frameworks and access to new data infrastructures in order to participate in and benefit from the ongoing AI boom. Prima facie, data-intensive AI approaches, especially artificial neural networks (ANNs) but also other approaches of machine learning (ML), increasingly enable and support the production of knowledge across all disciplines.

To conceptualize these shifts as mere effects of technology, however, arguably oversimplifies the interrelationship between technology and society. Since its learning capabilities rely on datafication, AI-based research is always connected to society from the start: AI uses its data to classify, categorize, and cluster society. What is more, as modern societies come to increasingly rely on scientific knowledge, any change in scientific practices and research methods brought about by AI technologies is bound to affect society at large (DFG 2020; Zhang et al. 2021). Research on the influence of AI on science is therefore of the utmost importance in order to comprehensively understand the effects of AI on present and future society. The term "science" here refers to natural sciences, social sciences, and humanities, and "scientific research" to those research methodologies within the sciences which use empirical and quantifiable data.

One of the big challenges is to understand and conceptualize the relationship between new technologies, specifically AI, and the epistemologies they enable: Are AI-based methods basically more efficient tools which continue non-AI methods, merely extending them in terms of velocity and scope? Or do they allow research in new methodological ways, to ask and answer novel questions? Or both on different levels? Which socio-political implications does data-driven research with AI entail, compared to other long-standing data-based research practices? Is AI-enabled data science still confined to numerical and quantifiable problems, or does it also give access to qualitative problems (e.g., problems of fuzziness; Seising, 2009)? How do AI-based methods try to reduce the dependency on (big) data, e.g., by making use of pre-trained models, or specific approaches like transfer learning or one-shot learning (Duan et al. 2017; Weiss/Khoshgoftaar/Wang 2016)?

To answer such questions, it is insufficient to merely look at AI models in a narrower sense (i.e., "learning algorithms") and deduce their impact. Since scientific research always occurs within specific *epistemic cultures* (Knorr-Cetina 1999; see also Fleck 1980 [1935]; Latour/Woolgar 1979), the impact of AI-based methods on scientific research is only determinable by closely observing the interplay of technology and research practices. Their alleged opacity is an obstacle here: While AI-based approaches make it possible to process data in novel ways (i.e., identifying, classifying, categorizing), they partially or completely disable researchers' abilities to comprehend and track these AI-based processes (Adadi/Berrada 2018; Sudmann 2019). In light of these observations, we follow the hypothesis that data-driven and AI-based methods enable new epistemologies precisely by transforming one of the most long-standing scientific practices of all: *AI changes the way researchers interact with and relate to data.* Hence, we believe it is important to compare the specific impact of ANNs and ML procedures with existing findings on digital methods in the sciences, most notably simulation, big data, and statistical probability (Ash/Kitchin/Leszczynski 2019; Gramelsberger 2011; Leonelli 2016; Krüger/Daston/Heidelberger 1987).

Therefore, we currently investigate how AI-based methods are situated in concrete and specific research environments which draw together technologies and practices. We carefully develop our findings from firsthand knowledge of outstanding current AI research projects operating within the novel conditions of data infrastructures. We also take into account the history of the specific methods in question, including their affordances, and contextualize these techno-practical configurations within an in-depth history of data-based scientific practices. It is intended that our observations will eventually inform new research approaches, as our findings will be fed back into the development of an AI-based system that structures and comprehends scientific content from several modalities including text, speeches, and meetings. This includes a multitude of components such as automatic speech recognition, segmentation to automatically divide the content into coherent chapters, and text summarization. With this system, we hope to support the answering of meta-research questions of AI as part of our project.

To this end, and given the prevalence of data-driven research across disciplines, we have set-up a *transdisciplinary research project* combining the expertise of three disciplines: thinking through the complex entanglements of technologies, culture, and practices is one of the core assets of *media studies*; providing an in-depth history of data and modeling practices in various methodological traditions is the key contribution of the *history of science*; developing a profound mathematical understanding of AI models and using computational methods to engineer a cutting-edge tool for using AI to study AI is the current task of our project in computer science. We combine the expertise of these three disciplines to study the socio-technical uses of AI in three carefully selected external research projects. From an original sample of close to 150 research projects working with AI-based methods in Europe ("How is Artificial Intelligence Changing Science?" 2023), we have chosen three projects from three different disciplines as the center of our investigation (film studies, sociology and climatology/Earth sciences). To capture the current changes brought about by AI in general and ML and ANNs in particular, our working groups in *media studies* (MS), *history of science* (HS), and *computer science* (CS) will combine the specific strengths of the most advanced methods from their respective fields. The project will allow a unique documentation and investigation in this pivotal decade. Otherwise, many of the traces of this historical shift which are obtainable now, will be irretrievably lost.

In the following, we have compiled a selection of theses that address some of the central aspects and considerations of our research group while also illustrating the range of different disciplinary perspectives on the various dimensions of AI in science. Each thesis is preceded by a quote, pointing to a larger topic to be further investigated in the course of the research project.

Thesis I: AI revolution

> Major economies are on the 'cusp of an AI revolution' that could trigger job losses in skilled professions such as law, medicine and finance, according to an influential international organisation. (Milmo 2023)

One obvious rhetoric in the discourse of so-called "AI" is that we are dealing with a "revolution". This seems to imply that some fundamental change will occur with the advent of AI technologies. This is easier said than understood. Often the rhetoric of "revolution" is simply used for marketing – because a new product, be it an advanced toothbrush or a new type of AI software, sells better if it is claimed to be a brand-new breakthrough of some kind. In that case "revolution" is used as a synonym to the entrepreneurial buzzword "disruption". It is normally not meant that a fundamental societal upheaval is to be expected (and that is the idea connected with the word "revolution" in the twentieth century), but just that there is a new product that displaces other products on the market. Basically, this is also the meaning of the recent and somewhat disturbing announcements of AI-producing companies that their own products might put humanity in danger – and call for regulation. If they are in fact so dangerous, why don't they simply stop producing these programs? It is more likely that they want to direct attention to how powerful their brand-new products might be or that the established players want to impede competition.

Besides that and if you take the claim of "revolution" more seriously, it is often not very clear what exactly is meant by that. If we take the example above, mentioned in *The Guardian*, a "revolution" was indeed the case if the job losses triggered by AI would lead to the fundamental impossibility of our (capitalist) societies, based on wage labor, to reproduce itself (on the following, see Schröter 2019). This problem of "technological unemployment" is actually an effect that was predicted by certain strands of Marxian theory already long ago, and long before AI. For many authors, this means that capitalism has to be overcome, or at least that radical political solutions, like unconditional basic income, have to be sought for. But is this really meant by the headline of *The Guardian*? Even if *The Guardian* is left-leaning, it can be doubted that it really wants to say that the (often postponed) terminal crisis of capitalism is now really here – with AI. That would be a "revolution" indeed. It is more likely that the article wants to say that certain professions that seemed safe so far are now also under the threat of automation. Although this might be bad for the people involved, this is nothing new. Many technological transformations happened in the last 150 years, many people lost their jobs, but also many new professions appeared. In the current situation, one presumably new development is the destabilization of the position of 'knowledge workers' and creative workers, i.e., subjects whose tasks have up till now not been automatable to a satisfying degree. If this shift will evoke a fundamental change in the dynamics of the division of labor remains to be a point of investigation. However, drawing on a wider perspective it can be said that certain rhetoric and discursive figures always return with new technologies (see Kümmel/Scholz/Schumacher 2004). New technologies are very often accompanied by utopian and dystopian ideas regarding their possible effects. To name only one example related to AI, that has also been the case with the internet (Schröter 2004). In the end, neither the worst fears came true nor the utopian paradise started. No (technologically driven) revolution happened, but given social structures were extended, accelerated and thereby transformed (but not in an abrupt "revolutionary" way). The internet is a very good example for that: Instead of leading immediately (dystopian) to a totalitarian hive-mind or (utopian) to a "frictionless capitalism" (a term coined by Bill Gates, see Schröter 2012) or even to a post-capitalist society, it became integrated into real existing capitalism (quite full of frictions), extending (to "friendship", into every moment and place of life etc.) and accelerating it – step by step. There is no reason to expect that this will be fundamentally different with AI. Neither the dystopian (AI will take over the world and kill everyone, capitalism collapses and this leads to total social disaster, etc.), nor the utopian (AI will solve all problems, a wonderful post-capitalist society will be born, etc.) visions will come true – but as always some of the good and some of the bad prospects will be realized and a lot more things will happen which were not expected or predicted at all. But that they were not expected doesn't amount to a "revolution" – that's just what history is.

If we now turn to our project called "How is Artificial Intelligence Changing Science?" – can we say that there is a kind of "scientific revolution" caused by AI (that means today mainly machine learning)? Given the state of the research we have done, it is too early to give a clear-cut answer – but our preliminary research shows that it might make sense to be cautious here too. On the basis of our research on how machine learning (and computer simulation) is used in high-energy physics (Schröter 2021; Radovic et al. 2018), we were able to test a claim regarding an alleged "scientific revolution" in science – the case is formulated in Anderson's (2008) much discussed paper on the "end of theory". His argument does not address AI directly – but the role of detecting patterns in large amounts of data which is exactly the task of many AI-systems today. He argues that the classical procedure of the natural sciences is now obsolete: While it was, prior to "big data", necessary to formulate a theory which then has to be tested in experiments, now it is enough to observe patterns and correlations in data. Theory is not needed anymore – a "scientific revolution" indeed. But at least for the case studied, this argument turned out to be wrong. Theory, very complicated theory, is still needed in particle physics. It predicts effects. Based on the theory, simulation models are generated that show how the patterns of the predicted effects would "look" like in the particle accelerator that is used to conduct experiments. The machine learning systems are trained with these simulated patterns – and then they filter out possible fitting patterns from the gigantic data stream produced by the accelerator. In this way, the predicted Higgs boson was found in 2012. Although the basic epistemology seems unchanged, in detail there are differences:

> The traditional way to analyze, or generate simulated, data is to first develop algorithms based on domain knowledge, then implement them in software, and use the resulting programs to analyze or generate data. This process is labor intensive, and analyzing complex datasets with many input variables becomes increasingly difficult and sometimes intractable. Artificial intelligence (AI) and the subfield of machine learning (ML) attack these problems in a different way: instead of humans developing highly specialized algorithms, computers learn from data how to analyze complex data and produce the desired results. There is no need to explicitly program the computers. Instead, ML algorithms use (often large amounts of) data to build models with relatively small human intervention. These models can then be applied to predict the behavior of new, previously unseen data, to detect anomalies or to generate simulated data. (Bourilkov 2019: 1f.)

That means: The application of AI systems leads to continuities and discontinuities at the same time. Our thesis is also: "Revolution" is a too narrow concept to describe the coexistence of continuities and discontinuities in the process of the diffusion of AI. One needs more differentiated concepts from media historiography (Schröter/Schwering 2014) to describe the effects of AI, even when only focused on the use of AI in different fields of science.

It is, of course, not necessary to conclude that the application of machine learning in other scientific fields follows the same trajectory. While in physics it seems to change nothing on a fundamental epistemic level (except for making use of far larger datasets as before), this might turn out to be different in other disciplines. This is essentially what our project tries to find out.

Thesis II: AI embedded

> [I]n order to grasp […] unwillingness to comply with mechanical innovation, we need to widen our perspective beyond machine technology. (von Oertzen 2017: 131)

For research on the implementation or the non-implementation of new technologies in various (scientific) fields, close attention should be paid to the infrastructures in which these technologies like machine learning are embedded. This allows for an analysis which refrains from conceptualizing abstract "technological enhancements" as the sole driving force of history. In this regard, non-implementations of AI technologies in scientific fields should not a priori be regarded as motivated by irrational conservatism or technophobia but must be researched within their specific political economy. Three examples from different scientific fields will serve to underpin this argument:

In the article quoted above, it is illustrated how the 19th-century Prussian census system by relying on "manual concepts, technologies, and practices of data power" (von Oertzen 2017: 129) managed to reach a similar level of effectiveness compared to other states which had implemented the Hollerith machine. Even after the eventual switch to machine-readable punch cards in the late 19th century which "enabled statisticians to accomplish tasks that were impossible to perform manually […] they rested firmly on the concepts and paper tools developed for manual use." (ibid.: 132) This episode can serve as an incentive to investigate thoroughly the conditions of the possibility of implementing new technologies instead of focusing on tech companies' accelerating announcements of revolutions and breakthroughs in their technological products.

Another important aspect of institutions' rigidity or refusal to implement new technologies has to do with what STS scholars Sheila Jasanoff, Ulrike Felt and others have coined as "sociotechnical imaginaries", i.e., "collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology." (Jasanoff 2015: 4). An episode from Western German computer history can serve as an example thereof: When Remington Rand delivered Europe's first large-scale computer system UNIVAC I to Frankfurt am Main in 1956, it caused considerable sensation about this "electronic brain". However, the plan to rent the machine to local companies for computational work ultimately failed and the UNIVAC I was shut down. As computer historian Corinna Schlombs argues, this failure was mainly due to Remington Rand's ignorance of "local customs and traditions" (Schlombs 2010: 98). This concerned the different organizations and sizes of German companies (139), European labor law (97), different infrastructural conditions, like electrical plugs (140) as well as the – proclaimed by a company report – "German users tend[ing] to be somewhat skeptical of the large scale systems" (ibid.: 139). Only after adapting to European conditions were Remington Rand's products able to gain a foothold in the European market. The story of the UNIVAC I points to the important question of why a technology does not fit into an environment, its types of organization, its infrastructures, and its sociotechnical imaginaries and which adaptive measures are taken to enable an implementation.

For the case of AI technology, such a focus on the infrastructural conditions of possibility (or impossibility) can further be helpful to grasp the different transformative speeds of human (scientific) practice. Instead of standing in awe of the daily releases and presentations and new models, close attention should be paid to a possible gap between modeling and implementation as to refrain from writing a mere history of ideas. As Urvi Sonawane and Matthieu Komorowski show in their contribution to this volume for the field of medical intensive care, "there is an increasing number of AI prototypes and early models being developed and trialled" (Sonawane/Komorowski 2023: 161). At the same time, "there seems to be a disproportionate disparity when it comes to translating these AI models from production to clinical evaluation." (ibid.: 161). Although the number of AI models released for the use in intensive care has risen significantly, implementation remains scarce. According to the authors, this is because, "the successful algorithms are less suited to be rolled out on a large-scale healthcare service or even across a country" (ibid.: 164) as well as the fact that "AI systems are notoriously difficult to integrate within and between systems" (ibid.: 164). Again, here it is the "problem" of different organizational systems and infrastructures being grown over a considerable amount of time, which yet complicates the AI models' widespread use in intensive care. This is of course not to foreclose the possibility that these systems can be fundamentally altered by the advent of new technologies. However, close attention should be paid to these different speeds of practical transformation by (AI) technology which can be addressed by an approach as outlined above.

Thesis III: Epistemological potentials

Statistics is the study of uncertainty. (Lindley 2000: 301)

Machine learning is essentially a form of statistics, and AI applications clearly display a statistical anatomy (Alpaydin 2016: 27). Hence one might think that the main task of machine learning approaches of AI would be the study of uncertainty. Indeed, specifically, artificial neural networks (ANNs) allow a new technical level of dealing with problems of uncertainty, for example dealing with incomplete information or predicting future events. Nevertheless, it would be insufficient to simply describe the general capabilities of ANNs in the processing of uncertainty. As a predictive technology, ANNs are, of course, in some way always related to problems of uncertainty, yet this does not accurately describe their enormous epistemic-technical capability to deal with different forms of vagueness or fuzziness related to visual or acoustic challenges of pattern recognition.

Over the course of the 19th century, statistical methods and probabilistic approaches took a successful hold in sciences as diverse as psychology and paleontology, sociology and astronomy, evolutionary biology and economic reasoning about risk and crime, in insurance and gambling. This unanimous shift towards quantitative methods came at a price. At its core lay the acceptance of less precision – or the new form of evidence which was later deemed as probabilistic revolution (Krüger/Daston/Heidelberger 1987). Back then, questions of fuzziness and uncertainty were intensively discussed by scientists such as Gustav T. Fechner and Pierre-Simon Laplace, as a shift away from the ideal of determinism that still prevailed at the beginning of that century. Currently, the boom in statistical AI in the form of ANNs, among others, makes the discussion of questions of uncertainty and fuzziness seem particularly urgent.

Already Claude Shannon's mathematical theory of communication formulated as a theorem on what digital technology in the form of computers was soon to achieve, namely dealing with problems of uncertainty, be it in relation to communication and its encryption or decryption or in relation to the prediction of flying objects in the application of radar technology. What corresponding communications technology or, ultimately, a computer achieved in the one case as well as in the other, in information-theoretical terms, is to distinguish between information and noise, and to make this possible as an exact calculation. And it was precisely for this purpose that the principle of binary circuitry proved to be particularly effective.

However, as became apparent in the course of the second half of the 20th century, certain more sophisticated problems of fuzziness were very difficult to solve for decades, e.g., enabling a computer to visually perceive its environment and objects in it. No matter which AI methods were used to approach such tasks, whether with so-called symbolic, rule-based AI or with subsymbolic AI in the form of ANNs or with approaches of so-called fuzzy sets, in the end, all these methods, despite selective progress, remained quite far away from what current AI implementations are capable of, until the 2000s.

It was only about 15 years ago that the situation changed significantly, as the important fields of AI work, computer vision or natural language processing, exemplify. Only then were computers able to cope with technically more demanding problems of fuzziness of various types much better.

We would like to briefly highlight this epistemic potential once again: Machine learning methods in the form of ANNs are in any case not only able to recognize patterns in complex data that are difficult for us humans to recognize due to their size and complexity and therefore present themselves to us as fuzzy, but which are themselves fuzzy and/or incomplete as statistical patterns. ANNs can produce usable output despite incomplete data or on the basis of fuzzy patterns, and they can do this by calculating not exactly, but approximately, i.e., quasi-fuzzy themselves. Thus, already the iterative, optimisation-oriented training process of machine learning methods can ideally be understood as a process of successive reduction of the prediction error, thereby approximating the real data distribution. Moreover, in the context of ANNs, there is now a broad portfolio of methods for dealing with uncertainty problems including ensemble methods, data augmentation methods, dropout and transfer learning.

Thesis IV: Big tech and academia

> One important feature of AI's modern R&D trajectory is that private companies native to the digital economy such as Google and Facebook are playing an increasingly important role in basic research activities that used to be the domain of academia. (Jurowetzki et al. 2021: 3)

As in many other fields, the conditions of possibility for applying AI processes in the sciences are shaped more than ever by the big tech industry.

First, there is an industry-wide tendency to provide universal, i.e., non-domain-specific, infrastructural support and tools to users. This includes systems such as ChatGPT.

Second, the big tech industry as well as leading AI companies such as Open AI or Antrophic are selectively engaged in solving fundamental problems of science in different domains for which a) technical-epistemic approaches of AI are particularly suitable and which b) should also redeem the claim to be socially responsible AI. In this respect, it cannot be surprising that some of the most important AI developments in this regard have taken place in the field of medicine.

Third, and this seems to us to be a particularly revealing area, the big tech industry is also the addressee for requests for support or funding of scientific projects that are developed by universities or private research institutions.

Fourth, it can be assumed that the big tech industry will drive the development of domain-specific tools and infrastructure offerings even more strongly in the future. Meta's scientific language model Galactica (Taylor et al. 2022) and Google's language model for medicine dubbed Med-PaLM (Singhal et al. 2023) are indicative of this trend.

The points listed here are certainly not specific to the big tech industry. On the contrary, it can be assumed that corresponding activities are generally driven by the tech industry, including start-ups. However, the more successful the respective activities and developments are, the greater the likelihood that either corresponding start-ups will be acquired or the big tech industry will develop similar tools or offerings, even if this potentially leads to legal conflicts.

From our perspective, at any rate, there is a serious transnational dependence of the sciences on industry, the precise conditions of which, in turn, urgently need to be researched on an interdisciplinary basis.

There are already signs that universities in Europe want to strengthen their independence from the big tech industry with regard to their research activities, but also in terms of teaching, while at the same time and to a certain extent paradoxically there is a university policy interest in promoting alliances between science and industry in the development of AI projects, whereby local and regional funding aspects may be of importance here.

One of the problems concerning the relationship between industry and science also includes the fact that, on the one hand, the industry generally has an interest in ensuring that universities are able to train sufficiently qualified scientists, while at the same time, it also has a considerable share in the fact that particularly qualified scientists leave universities for industry and, in part, industry is increasingly moving to promote the internal training of IT specialists.

The aspects mentioned above thus concern fundamental questions about the conditions of digital science in the present. It seems important to us, however, that all the industrial and infrastructural problems and challenges mentioned above are causally linked to the specific potentials of subsymbolic AI.

The dependencies indicated here must be critically questioned, not only as a questionable contrast between a big tech industry worthy of criticism and university research supposedly independent from the outset, but in general with a view to the possibility of sustainably protecting technology from misuse, whether in the context of the private sector or with a view to public/state structures. Democratic states may lose their democratic or progressive status, much as the policies of large corporations may change drastically, and with it the question of what purposes AI is used for in the first place. Based on our observations so far, having to seriously consider the long-term consequences of AI is a relatively new phenomenon. Until recently, it seemed important to free AI from speculative discourse, and rather address problems of AI's present, but in light of recent developments, it does indeed seem necessary to extrapolate current developments and their speed to what problems will arise not just now, but in 5 or 10 years.

Thesis V: Expert crisis

> AI experts are in short supply. That's making the skills crisis worse. (Headline of an article on ZDnet, Hughes 2022)

One consequence of the fact that approaches of ANNs and other forms of machine learning could not really be used comprehensively or for advanced tasks, e.g., in many fields of science, is that a corresponding tradition of expertise was missing at the beginning of the current AI boom. Only a relatively small number of researchers in Europe and the US focused on or worked with such approaches in typical fields of AI research before 2016. Moreover, those who worked with ANNs in computer science in the 1990s and early 2000s, for example, had significant problems themselves at the time in being able to finance and publish their research at all. The establishment of the term 'deep learning' for ANNs from around 2006 onwards had a lot to do with the reputation of ANNs as ultimately being more or less a dead end in AI research, or as ultimately not being a target-oriented approach for many advanced problems in computer science. Accordingly, for a long time, it was an internationally very manageable community that continued unperturbed with ANN-based AI, in German-speaking countries, for example, research groups around Helge Ritter, Jürgen Schmidhuber and Alexander Waibel.

Against this background, the contrast to today's situation could hardly be more extreme. Especially from 2016 on, i.e., since the success of AlphaGo, a gold-rush mood has developed rapidly, which in turn not only affected the scientific field alone but more or less the society as a whole.[1]

From a disciplinary point of view, it is obvious that computer science in particular has benefited from the corresponding AI boom. In fact, it has not hurt the discipline in principle to have underestimated the epistemic potential of ANNs at the time.

Computer science is the big winner of the current AI boom in two respects: on the one hand, because of its historical core competence with respect to both the development and critical reflection of AI, and on the other hand, because of its now once again strengthened role as a collaborative or auxiliary discipline of other subjects. It may be that some disciplines, such as mathematics or physics, are not dependent on the external competencies of computer science to develop AI models for their purposes, but the humanities, cultural studies and social sciences are (even though in these areas knowledge on computer science has increased significantly in the recent past).

At the same time, the cultural sciences, social sciences, and humanities also benefit prima facie considerably from the sustained boom in AI. This applies, among other things, to philosophy, whose expertise has been called upon for some years now, especially for ethical issues in AI.

Finally, this concerns the interdisciplinary and transdisciplinary research field of digital humanities. Even if the corresponding orders of magnitude are difficult to estimate, one can certainly argue that AI, and especially forms of generative AI, have an important catalytic function in significantly expanding

1 A German platform currently lists 152 institutes and other institutions at German universities that conduct research with/on AI (Lernende Systeme 2022). Also see Huber/Huth/Alsabah (2020), a Bitkom survey according to which there are about 220 AI professorships in Germany at the time of the survey. Finally, one could point to the 100 AI Professorships Initiative, initiated in 2018 (BMBF 2022).

the possible uses of computers and Big Data, in the humanities and cultural studies, whether in research, or also in other areas such as teaching.

However, on the basis of our studies so far, we can state that certain subjects or research fields, such as climatology, are only gradually incorporating AI approaches to their research questions, and to a rather limited extent, and in some cases, there is also a rather great skepticism, perhaps even a certain conservatism, about using corresponding technologies.

Thesis VI: Sociological split seconds

> Computational social science is an interdisciplinary field that advances theories of human behavior by applying computational techniques to large datasets from social media sites, the Internet, or other digitized archives such as administrative records. Our definition forefronts sociological theory because we believe the future of the field within sociology depends not only on novel data sources and methods, but also on its capacity to produce new theories of human behavior or elaborate on existing explanations of the social world. (Edelmann et al. 2020)

A general phenomenon in the research landscape is the bifurcation of disciplines in a general and a computational branch. For instance, these twin disciplines have become a reality in many fields of the social sciences and are even traceable by citation analysis. Is computational sociology, which emerged towards the end of the 20th century, out of tune with the offline society or sociological theory? Will such disciplines eventually divide for good despite the integrative gestures the computational disciplines may provide? It is entirely conceivable that digital historians will seize to travel to the communities' main conferences (like the "Historikertag"). It is possible that computational Earth scientists find it easier to talk method with colleagues from the digital humanities than with colleagues returning from the field with samples and earth on their boots.

Technological advances provide new tools. Arguably, these provide new gravitational forces towards specific scientific methods and topics. The integration of new kinds of data practices across disciplines is all but new. Throughout the last centuries, the sciences, social sciences and humanities have benefited significantly from the availability of data. The empirical sciences of the 17th century, the social sciences of the 19th century and the digital humanities of the 21st century all profited from the influx of serialized and quantified types of information into their research methods. While it is probably the humanities which will be affected most by AI's conquest of the qualitative dimension, the social sciences present a puzzle.

For instance, the emergence of a separate field of computational sociology is a most interesting case, given the high affinity of sociology to data and empirical methods. This discipline can be said to have co-emerged with census taking and the intensified collection of data about the social from Adolphe Quetelet onwards. His "social physics" and his training with leading astronomers in Paris did not only lead him to stipulate neutral laws of the social, but let him develop moral statistics, a brand of criminology if not surveillance.

Given this high involvement of the discipline with social data, the division into separated disciplines of sociology may come as a surprise. This is not to say that sociologists do not embrace new data technologies. Attempts to map all articles in the Web of Science according to their level of AI-related methods show the social sciences almost as open to AI applications as the physical sciences or the life sciences (Gargiulo et al. 2022, fig. 1). Yet, this adaptability does not seem to appeal to all sociologists and at the moment several fields of knowledge split into computational versions of themselves. While this may prove to be a passing occurrence, it can also be an indication that AI-based methodologies are not perceived as empirical in the traditional sense.

Thesis VII: Data colonialism

> [...] I wonder whether data colonialism goes far enough to prompt a decolonial shift in thinking, assuming again that we are in the realm of Quijano and the modernity/coloniality school. Because the concept is more concerned with datafication as resource extraction, and seems less concerned with the key decolonial insight that Europe convinced itself and others that it has a privileged objective position from which it may make universal assertions and claims. (Mumford 2022: 1512)

In their widely discussed studies on data colonialism Nick Couldry and Ulisses Mejias (2019a; 2019b) claim that a new regime of data extraction has emerged. They see a logic of colonial dispossession at work because most data are collected for free. After all, this annexation is happening on a global scale, and results in huge profits for a very small group of people. Most AI applications still hinge on the availability of vast datasets and are part and parcel of the said colonialism. The scope of this new raw material is lamented and the problematic uses of a global surveillance capitalism are evoked. Some emphasize the correlation of systems of data extraction with systems of value (Thatcher/O'Sullivan/Mahmoudi 2016; Gray 2023). According to Denusa Mumford (2022), it seems questionable if the term colonization is already put to its best use in this discussion. The critiques mainly address capitalist strategies, especially the dynamics of primitive accumulation. While this dynamic is involved with the Global North as colonizer, the term points towards the fate of the Global South, but fails to show any specific engagement. Recent calls for epistemic decolonization remain unheard. A decolonial approach would entail efforts to decentralize one's position, to seek out "other worldings", to include specifics of the discussions from these regions, and to acknowledge the fundamental diversity of approaches. Thus, the diagnosis of data colonialism might not go far enough yet and would benefit from a deeper involvement with divergent perspectives from the outside, from the effects of data collection at the margins.

The history of data tries to make inroads and add to a fuller picture of the specific performative effects of data collection at various points of a data journey (Aronova/von Oertzen/Sepkoski 2017; Leonelli/Tempini 2020). The global data infrastructure is clearly built on an overexposure of marginalized and colonized bodies to various kinds of metrics (Hacking 1986; Lemov 2015; Radin 2017; TallBear 2013). Even these early colonial statistics and data collections already had radical effects. Indigenous communities could suddenly be shown to "go extinct" or dwindle under the curse of hunting parties of colonialists, hard physical labor and new contagious diseases (Rowse 2017; Malègue 2018; Renard 2021). The most widely discussed case of colonial statistics though is the performative effect and deep influences of the process of counting itself (Zimmerman 1999; Schlicht/Ledebur/Echterhölter 2021). The famous case in point is probably the Indian census, where the British used the categories of "caste" for the enumeration of all Hindus. Although castes existed in pre-colonial times, their statistical versions made them more rigid, scientifically defined, prominent, and publicly contested or lobbied for (Cohn 1987; Appadurai 1994; Dirks 2001). This shows that the stakes are high for any category, classificatory scheme, or label used on social data. The history of data classification is but one aspect of the decolonization of our rapidly growing data architectures.

During German colonialism, data collections and land surveys consistently relied on European legal notions and hence implemented foreign protocols.

To reverse this and similar processes, several initiatives are trying to arrive at more fitting frameworks for data collection in non-industrialized societies today (Abdilla 2021). To involve communities and to encourage participation is but one strategy to improve a technology that is perceived as "White" (Cave/Dihal 2020). Against this background, current initiatives for "indigenous data sovereignty" gain importance (Santos 2018; Kukutai/Taylor 2019; Lewis 2020), and fluid identities are being discussed as a blueprint to build better modeled and networked data infrastructures (Chun et al. 2019). Do indigenous perceptions of the non-human help to "breed" better algorithms? Who decides upon the categories used in clustering or classification in which region of the world? What would it mean, with Denusa Mumford, to arrive at decolonial data architectures?

Thesis VIII: The labor landscape shift

> ChatGPT and the like do improvise, promising to destabilize a lot of white-collar work, regardless of whether they eliminate jobs or not. (Lowrey 2023)

In stark contrast to our long-standing expectations and to great surprise, AI is not automating routine tasks or physical work as its first official act. Instead, generative AI impacts highly-skilled creative and knowledge workers by producing creative and knowledgeable output. In this respect, AI sets itself apart from previous technological developments.

This new realization is fueled by several contemporary developments in AI research. First and foremost, there has been remarkable progress in the area of language models leading to singular models capable of fulfilling diverse tasks such as creative content generation, summarization, translation, and code generation – to name a few of them. By defining instructions and prompts, many more tasks are conceivable that previously required meaningful investments and specialized systems. There are already several applications that give a first glimpse of the potential for and the impact on certain professions. Software engineers are fast to adopt new technical tools, such that it is no surprise that GitHub Copilot, a coding assistant, already amassed over a million users and is "behind an average of 46% of a developers' code" (Zhao 2023).

There are similar applications for other areas such as SciSpace Copilot, aiming to help with scientific literature research, or Casetext, which assists with legal research. For now, these systems are meant to be assistant to workers rather than to act in full autonomy. However, AI research is in rapid development[2]. The GitHub CEO, Thomas Dohmke, for example, claims that GitHub CoPilot will "sooner than later" (Scheffler 2023) write "80% of the code" (ibid.). It is an open question if this massive productivity gain will result in job losses but at the very least it reshapes the nature of some professions, adding oversight of and delegation to LLMs as a major bullet point to many job descriptions. Notably, Meta itself, as one of the leaders in AI research, announced to reduce hiring meaningfully going forward and focus instead on developer productivity, with AI toolings like code assistants and chatbots being a major part of that equation.

A relevant study in this context was conducted by Eloundou et al. (2023) in which the authors try to identify jobs that are exposed to LLM technology. While they neither provide a timeline nor make predictions on the impact on the labor market, it is found that higher-income occupations are more affected. Mathematicians, writers and authors, tax preparers, legal secretaries, or proofreaders are among a set of professions that are fully or close to fully exposed to LLMs; that means that LLMs take up to 100% of their occupational activities.[3]

At the same time, there is another area of AI research swiftly progressing and bringing completely different qualities to the debate. With the rise of recent diffusion models such as Adobe Firefly, Unity Muse, Midjourney, and Stable Diffusion, AI image synthesis is widely popularized, from the generation of digital art to photorealistic art. Point-E or Builder Bot are first approaches to 3D content generation while Imagen Video and Make-A-Video conceptualize video synthesis. Remarkably, the public outcry about this line of research has been much louder and more popular, leading to several copyright lawsuits and massive fears about potential job replacement. Presumably, that is, because visual output is more tangible and inaccuracies are not as crucial or obvious. Exemplary, Hollywood is one of several epicenters of this debate. While Disney just released the TV show *Secret Invasion* with an AI-generated opening, the actors guild SAG-AFTRA went on strike with one major contentious point being

2 Also see the following thesis, "AI's Self-Evolution".

3 The positive news for us in the research community is that professions related to science and critical thinking were found to have a low exposure to LLMs.

the digital replication of actors in films and shows. Again, the impacts on the labor market are hard to predict, however, in a *New York Times* (Roose 2022) article artists anecdotally describe the transformation of their work with the arrival of AI tools, while a report by *Rest of World* (Zhou 2023) already claims an effect of AI on the job market of game illustrators.

A recent McKinsey study (Chui et al. 2023) predicts a massive transformation of our economy and the labor market as a result of generative AI, encompassing both large language models and diffusion models. The report claims "that half of today's work activities could be automated" (ibid.). That is, however, on a largely uncertain timeline "between 2030 and 2060" (ibid.). In accordance with the study by Eloundou et al. and our own assumptions, the report eventually sees high-income workers as the most impacted group of this transformation. And while no one dares to make a definitive prediction on the impact of the labor market, "it's important to be honest that it's increasingly going to make some jobs not very relevant" (Altman 2022) as Sam Altman himself puts it in an article on his website. At the very least, AI is predestined to alter the labor landscape profoundly.

Thesis IX: AI's self-evolution

> [A] Large Language Model (LLM) is capable of improving its performance [...] by training on its own generated labels. (Huang et al. 2022)

The landscape of AI is undergoing a transformative shift as we witness the emergence of a cycle of self-evaluating and self-improving systems. Large language models are at the forefront of this development, demonstrating the ability to assess the quality of and improve upon their own generations. At the same time, generated output can be used to leverage and train the next generation of models or distill knowledge into more efficient ones allowing to deploy models on a large scale. Several recent developments are indicative of this development.

To begin with, current LLMs are cheaper and already produce comparable output compared to human labelers on crowdsourcing platforms such as Mechanical Turk in many tasks (Gilardi/Alizadeh/Kubli 2023). At the same time, studies like the one by Veselovsky, Horta and West (2023) found that crowd workers increasingly utilize LLMs themselves to complete their tasks. It is foreseeable that this will supercharge data collection in many cases, increasing the quantity of data while reducing the time needed.

Separately, there are a number of developments in AI research itself that suggest an increasingly fast cycle of self-evaluation and improvement. For example, the popular Alpaca model from Stanford (Taori et al. 2023) uses GPT-3 to produce training data to fine-tune and align Meta's LLaMA language model (Touvron et al. 2023). All this happened within two weeks of the release of LLaMA, further indicating the rapid speed at which improvements are achieved. Taori et al. use a recent trend in AI called "Self-Instruct" (Wang et al. 2023) to automatically generate training data where a number of seed tasks is defined and an LLM generates new instructions and corresponding instruction-answer pairs for them.

In addition, there are other ways to utilize current LLMs. Traditionally, LLMs but also other AI systems are evaluated by certain automated metrics. However, these metrics are generally not perfect and only offer correlation with human judgment to a certain degree. This prompted several researchers to develop performance metrics based on the judgment of LLMs, such as GPTScore (Fu et al. 2023) or GEMBA (Kocmi et al. 2023). It is also becoming more common to evaluate LLMs only relative to each other using rating mechanisms like ELO (see Chatbot Arena by Zheng et al. 2023). One pressing question that may arise now, is whether it is possible to use these capabilities in the training process of an LLM. Presently, reinforcement learning from human feedback (RLHF; Christiano et al. 2017) is used by state-of-the-art models like GPT-3. As part of this, a reward model is trained based on human feedback. This reward model can then be used as a proxy of human feedback while fine-tuning the LLM. It is conceivable though to use a fine-tuned LLM in place of humans to create synthetic ranking data for the training of reward models. This technique is further called reinforcement learning from AI feedback (RLAIF) and can potentially be applied in an iterative fashion (Bai et al. 2022).

It also becomes more prevalent to use these self-evaluation capabilities to improve the performance of current LLMs at inference time. Even without further training or fine-tuning, the model can reflect or critique itself by passing the output *again* into the model and refine its output (Gou et al. 2023; Shinn/Labash/Gopinath 2023; Xiao et al. 2023).

Now, with these trends reshaping AI, one could be tempted to speculate about a vicious self-improvement cycle leading to the "singularity". However, it is important to temper these advances with the recognition that current AI systems including LLMs remain narrow in nature and are far from achieving artificial general intelligence (AGI), thereby dispelling the notion of an imminent transformative "singularity" or emergence of "superintelligence" driven by existing AI paradigms. For this to happen, it might need a fundamentally different approach to AI as suggested by Yann LeCun (2022) and others.

List of references

Abdilla, Angie/Kelleher, Meghan/Shaw, Rick/Yunkaporta, Tyson (2021): Out of the Black Box: Indigenous Protocols for AI, Melbourne: Deakin University (https://hdl.handle.net/10536/DRO/DU:30159239).

Adadi, Amina/Berrada, Mohammed (2018): "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)." In: IEEE Access 6, pp. 52138–52160.

Alpaydin, Ethem (2016): Machine Learning: The New AI, Cambridge, MA: The MIT Press.

Altman, Sam (2022): "DALL•E 2.", April 6, 2022 (https://blog.samaltman.com/dall-star-e-2).

Anderson, Chris (2008): "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.", June 23, 2008 (https://www.wired.com/2008/06/pb-theory/).

Appadurai, Arjun (1996): "Number in the Colonial Administration." In: Arjun Appadurai (ed.), Modernity at Large. Cultural Dimensions of Globalization, Minneapolis, MN: University of Minnesota Press, pp. 114–139.

Aronova, Elena/von Oertzen, Christine/Sepkoski, David (2017): "Introduction: Historicizing Big Data." In: Osiris 32/1, pp. 1–17.

Ash, James/Kitchin, Rob/Leszczynski, Agnieszka (eds.) (2019): Digital Geographies. Los Angeles, CA. et al.: SAGE.

Bai, Yuntao/Kadavath, Saurav/Kundu, Sandipan/Askell, Amanda/Kernion, Jackson/Jones, Andy/Chen, Anna/et al. (2022): Constitutional AI: Harmlessness from AI Feedback, arXiv Preprint (https://doi.org/10.48550/arXiv.2212.08073).

BMBF (Bundesministerium für Bildung und Forschung) (2022): "Stark-Watzinger: 100. zusätzliche KI-Professur wurde besetzt.", May 6, 2022 (https://www.bmbf.de/bmbf/shareddocs/pressemitteilungen/de/2022/05/030522-KI-Professoren.html).

Bourilkov, Dimitri (2019): "Machine and Deep Learning Applications in Particle Physics". In: International Journal of Modern Physics A 34/35, 1930019.

Cave, Stephen/Dihal, Kanta (2020): "The Whiteness of AI." In: Philosophy and Technology 33, pp. 685–703.

Christiano, Paul F./Leike, Jan/Brown,Tom/Martic, Miljan/Legg, Shane/Amodei, Dario (2017): Deep Reinforcement Learning from Human Preferences, arXiv Preprint (https://doi.org/10.48550/arXiv.1706.03741).

Chui, Michael/Hazan Eric/Roberts Roger/Singla Alex/Smaje Kate/Sukharevsky Alex/Yee Lareina/Zemmel Rodney (2023): "The economic potential of generative AI: The next productivity frontier.", June 2023 (https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-AI-the-next-productivity-frontier).

Chun, Wendy/Apprich, Clemens/Steyerl, Hito/Cramer, Florian (2019): Pattern Discrimination, Minneapolis and London: University of Minnesota Press, and Lüneburg: meson press.

Cohn, Bernard S. (1987): "The Census, Social Structure, and Objectification in South Asia." In: Bernard S. Cohn (ed.), An Anthropologist among the Historians and Other Essays, New Delhi: Oxford University Press, pp. 224–254.

Couldry, Nick/Mejias, Ulises A. (2019a): The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism, Stanford, CA: Stanford University Press.

Couldry, Nick/Mejias, Ulises A. (2019b): "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject." In: Television & New Media 20/4, pp. 336–349.

Dastin, Jeffrey (2023): "Microsoft attracting users to its code-writing, generative AI software.", January 25, 2023 (https://www.reuters.com/technology/microsoft-attracting-users-its-code-writing-generative-ai-software-2023-01-25/).

Deutsche Forschungsgemeinschaft (2020): Digitaler Wandel in den Wissenschaften, Bonn: Deutsche Forschungsgemeinschaft (https://doi.org/10.5281/zenodo.4191345?).

Dirks, Nicholas B. (2001): Castes of Mind: Colonialism and the Making of Modern India, Princeton: Princeton University Press.

Duan, Yan/Andrychowicz, Marcin/Stadie, Bradly/Ho, Jonathan/Schneider, Jonas/Sutskever, Ilya/Abbeel, Pieter/Zaremba, Wojciech (2017): "One-Shot Imitation Learning." In: Advances in Neural Information Processing Systems (NIPS, 2017), Long Beach, CA, USA, pp. 1087–1098.

Edelmann, Achim/Wolff, Tom/Montagne, Danielle/Bail, Christopher A. (2020): “Computational Social Science and Sociology.” In: Annual Review of Sociology 46/1, pp. 61–81.

Eloundou, Tyna/Manning, Sam/Mishkin, Pamela/Rock, Daniel (2023): GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models, arXiv Preprint (https://doi.org/10.48550/arXiv.2303.10130).

Fleck, Ludwik (1980 [1935]): Entstehung und Entwicklung einer wissenschaftlichen Tatsache, Frankfurt a. M.: Suhrkamp.

Fu, Jinlan/Ng, See-Kiong/Jiang,Zhengbao/Liu, Pengfei (2023): GPTScore: Evaluate as You Desire, arXiv Preprint (https://doi.org/10.48550/arXiv.2302.04166).

Gargiulo, Floriana/Fontaine, Sylvain/Dubois, Michel/Tubaro, Paola (2022): A meso-scale cartography of the AI ecosystem, arXiv Preprint (https://doi.org/10.48550/arXiv.2212.12263).

Gilardi, Fabrizio/Alizadeh, Meysam/Kubli, Maël (2023): “ChatGPT Outperforms Crowd-Workers for Text-Annotation Tasks.” In: Proceedings of the National Academy of Sciences 120/30, e2305016120.

Gou, Zhibin/Shao, Zhihong/Gong, Yeyun/Shen, Yelong/Yang, Yujiu/Duan, Nan/Chen, Weizhu (2023): CRITIC: Large Language Models Can Self-Correct with Tool-Interactive Critiquing, arXiv Preprint (https://doi.org/10.48550/arXiv.2305.11738).

Gramelsberger, Gabriele (ed.) (2011): From Science to Computational Sciences Studies in the History of Computing and its Influence on Today‘s Sciences, Zürich: Diaphanes.

Gray, Catriona (2023): “More than Extraction: Rethinking Data’s Colonial Political Economy.” In: International Political Sociology 17/2, pp. 1–20.

Hacking, Ian (1986): “Making Up People.” In: Thomas L. Heller/Morton Sosna/David E. Wellbery (eds.), Reconstructing Individualism. Autonomy, Individuality, and the Self in Western Thought, Stanford, CA: Stanford University Press, pp. 222–236.

Huang, Jiaxin/Shane Gu, Shixiang /Hou, Le/Wu, Yuexin/Wang, Xuezhi/Yu, Hongkun/Han, Jiawei/et al. (2022): Large Language Models Can Self-Improve, arXiv Preprint (https://doi.org/10.48550/arXiv.2210.11610).

Huber, Marco/Huth, Michael/Alsabah, Nabil (2020): KI-Forschung in Deutschland – Der schwere Weg zu 100 neuen KI-Professuren, Berlin: Bitkom.

Hughes, Owen (2022): “AI Experts Are in Short Supply. That’s Making the Skills Crisis Worse.” May 20, 2022 (https://www.zdnet.com/article/ai-experts-are-in-short-supply-thats-making-the-skills-crisis-worse/).

Jasanoff, Sheila (2015): "Future Imperfect. Science, Technology, and the Imaginations of Modernity." In: Sheila Jasanoff/Sang-Hyun Kim (eds.), Dreamscapes of Modernity. Sociotechnical Imaginaries and the Fabrication of Power, Chicago and London: The University of Chicago Press, pp. 1–33.

Jurowetzki, Roman/Hain, Daniel/Mateos-Garcia, Juan/Stathoulopoulos, Konstantinos (2021): The Privatization of AI Research (-ers): Causes and Potential Consequences–From university-industry interaction to public research brain-drain?, arXiv Preprint (https://doi.org/10.48550/arXiv.2102.01648).

Knorr-Cetina, Karin (1999): Epistemic Cultures: How the Sciences Make Knowledge, Cambridge, MA: Harvard University Press.

Kocmi, Tom/Federmann, Christian (2023): Large Language Models Are State-of-the-Art Evaluators of Translation Quality, arXiv Preprint (https://doi.org/10.48550/arXiv.2302.14520).

Krüger, Lorenz/Daston, Lorraine J./Heidelberger, Michael (eds.) (1987): The Probabilistic Revolution, Cambridge, MA: The MIT Press.

Kukutai, Tahu/Taylor, John (eds.) (2016): Indigenous Data Sovereignty: Toward an Agenda, Canberra: Australian National University Press.

Kümmel, Albert/Scholz, Leander/Schumacher, Eckhardt (2004): Einführung in die Geschichte der Medien, Paderborn: Fink.

Latour, Bruno/Woolgar, Steve (1979): Laboratory Life. The Construction of Scientific Facts, Princeton, NJ: SAGE.

LeCun, Yann (2022): A Path Towards Autonomous Machine Intelligence, OpenReview Preprint (https://openreview.net/forum?id=BZ5a1r-kVsf).

Lemov, Rebecca (2015): Database of Dreams: The Lost Quest to Catalog Humanity, New Haven, CT: Yale University Press.

Leonelli, Sabina (2016): Data-Centric Biology. A Philosophical Study, Chicago and London: University of Chicago Press.

Leonelli, Sabina/Tempini, Niccolò (eds.) (2020): Data Journeys in the Sciences, Cham: Springer Nature.

Lernende Systeme (2023): "KI-Landkarte.", continuously updated (https://www.plattform-lernende-systeme.de/ki-landkarte.html).

Lewis, Jason E. (ed.) (2020): Indigenous Protocol and Artificial Intelligence Position Paper, Honolulu, HI: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (https://doi.org/10.11573/spectrum.library.concordia.ca.00986506).

Lindley, Dennis V. (2000): "The Philosophy of Statistics." In: Journal of the Royal Statistical Society: Series D (The Statistician) 49/3, pp. 293–337.

Lowrey, Annie (2023): "How ChatGPT Will Destabilize White-Collar Work.", January 20, 2023 (https://www.theatlantic.com/ideas/archive/2023/01/chatgpt-ai-economy-automation-jobs/672767/).

Malègue, Fanny (2018): "The Empire in a Census Table: Enumerating the Population and Governing the French Caribbean Colonies after the Seven Years' War." In: Histoire & Mesure 2, pp. 93–114.

Milmo, Dan (2023). "AI revolution puts skilled jobs at highest risk, OECD says.", July 11, 2023 (https://www.theguardian.com/technology/2023/jul/11/ai-revolution-puts-skilled-jobs-at-highest-risk-oecd-says).

Mumford, Denusa (2022): "Data Colonialism: Compelling and Useful, but Whither Epistemes?" In: Information, Communication & Society 25/10, pp. 1511–1516.

Radin, Joanna (2017): "'Digital Natives': How Medical and Indigenous Histories Matter for Big Data." In: Osiris 32, pp. 43–64.

Radovic, Alexander/Williams, Mike/Rousseau, David/Kagan, Michael/Bonacorsi, Daniele/Himmel, Alexander/Aurisano, Adam/Terao, Kazuhiro/Wongjirad, Taritree (2018): "Machine Learning at the Energy and Intensity Frontiers of Particle Physics." In: Nature 560, pp. 41–48.

Renard, Léa (2021): "Vergleichsverbot? Bevölkerungsstatistiken und die Frage der Vergleichbarkeit in den deutschen Kolonien (1885–1914)." In: Kölner Zeitschrift für Sozialpsychologie 73, pp. 169–194.

Roose, Kevin (2022): "A.I.-Generated Art Is Already Transforming Creative Work.", October 21, 2022 (https://www.nytimes.com/2022/10/21/technology/ai-generated-art-jobs-dall-e-2.html).

Rowse, Tim (2017): "The Statistical Table as Colonial Knowledge." In: Itinerario 41/1, pp. 51–73.

Santos, Boaventura de Sousa (2018): The End of the Cognitive Empire: The Coming of Age of Epistemologies of the South, Durham: Duke University Press.

Scheffler, Ian (2023): "GitHub CEO says Copilot will write 80% of code 'sooner than later'", June 17, 2023 (https://www.freethink.com/robots-ai/github-copilot).

Schlicht, Laurens/Ledebur, Sophie/Echterhölter, Anna (2021): "Data at the Doorstep. Intimacy and Formalization." In: Science in Context 34/4, pp. 1–18.

Schlombs, Corinna (2010): Productivity Machines: Transatlantic Transfers of Computing Technology and Culture in the Cold War, Ann Arbor, MI: UMI Dissertations Publishing.

Schröter, Jens (2004): Das Netz und die Virtuelle Realität. Zur Selbstprogrammierung der Gesellschaft durch die universelle Maschine, Bielefeld: transcript.

Schröter, Jens (2012): "The Internet and 'Frictionless Capitalism'." In: tripleC 10/2, pp. 302–312.

Schröter, Jens (2019): "Digitale Medientechnologien und das Verschwinden der Arbeit." In: Thomas Bächle/Caja Thimm (eds.), Mensch und Maschine: Freund oder Feind? Mensch und Technologie im digitalen Zeitalter, Wiesbaden: Springer, pp. 183–210.

Schröter, Jens (2021): "KI und die Wissenschaften. Das Beispiel der Teilchenphysik." In: Nicole Brandstetter/Ralph-Miklas Dobler/Daniel Jan Ittstein (eds.), Mensch und Künstliche Intelligenz. Herausforderungen für Kultur, Wirtschaft und Gesellschaft, München: UVK, pp. 51–58.

Schröter, Jens/Schwering, Gregor (2014): "Modelle des Medienwandels und der Mediengeschichtsschreibung." In: Jens Schröter (ed.), Handbuch Medienwissenschaft, Stuttgart and Weimar: Metzler, pp. 179–190.

Seising, Rudolf (ed.) (2009): Views on Fuzzy Sets and Systems from Different Perspectives: Philosophy and Logic, Criticisms and Applications, Berlin and Heidelberg: Springer (https://doi.org/10.1007/978-3-540-93802-6).

Shinn, Noah/Labash, Beck/Gopinath, Ashwin (2023): Reflexion: an Autonomous Agent with Dynamic Memory and Self-reflection, arXiv Preprint (https://doi.org/10.48550/arXiv.2303.11366).

Singhal, Karan/Azizi, Shekoofeh/Tu, Tao/Mahdavi, Sara/Wei, Jason/Won Chung, Hyung /Scales, Nathan/et al. (2023): "Large Language Models Encode Clinical Knowledge." In: Nature 620, pp. 172–180.

Sonawane, Urvi/Komorowski, Matthieu (2023): "Artificial Intelligence in Medicine. Potential Applications and Barriers to Deployment." In: Andreas Sudmann/Anna Echterhölter/Markus Ramsauer/Fabian Retkowski/Jens Schröter/Alexander Waibel (eds.), Beyond Quantity: Research with Subsymbolic AI, Bielefeld: transcript, pp. 155-178.

Sudmann, Andreas (2019): "On the Media-Political Dimension of Artificial Intelligence. Deep Learning as a Black Box and OpenAI." In: Andreas Sudmann (ed.), The Democratization of Artificial Intelligence. Net Politics in the Era of Learning Algorithms, Bielefeld: transcript. pp. 223–244.

Syed, Ali (2016): "A Brief Introduction to Decolonial Computing." In: Crossroads (Association for Computing Machinery) 22/4, pp. 16–21.

TallBear, Kimberly (2013): Native American DNA: Tribal Belonging and the False Promise of Genetic Science, Minneapolis, MN: University of Minnesota Press.

Taori, Rohan/Gulrajani, Ishaan/Zhang, Tianyi/Dubois, Yann/Guestrin, Carlos/Liang, Percy/Hashimoto, Tatsunori B. (2023): "Alpaca: A Strong, Replicable Instruction-Following Model.", March 13, 2023 (https://crfm.stanford.edu/2023/03/13/alpaca.html).

Taylor, Ross/Kardas, Marcin/Cucurull, Guillem/Scialom, Thomas/Hartshorn, Anthony/Saravia, Elvis/Poulton, Andrew/et al. (2022): Galactica: A Large Language Model for Science, arXiv Preprint (https://doi.org/10.48550/arXiv.2211.09085).

Thatcher, Jim/O'Sullivan, David/Mahmoudi, Dillon (2016): "Data Colonialism through Accumulation by Dispossession: New Metaphors for Daily Data." In: Environment and Planning D: Society and Space 34/6, pp. 990–1006.

Touvron, Hugo/Lavril, Thibaut/Izacard, Gautier/Martinet, Xavier/Lachaux, Marie-Anne/Lacroix, Timothée/Rozière, Baptiste/et al. (2023): LLaMA: Open and Efficient Foundation Language Models, arXiv Preprint (https://doi.org/10.48550/arXiv.2302.13971).

Veselovsky, Veniamin/Horta, Ribeiro Manoel/West, Robert (2023): Crowd Workers Widely Use Large Language Models for Text Production Tasks, arXiv Preprint (https://doi.org/10.48550/arXiv.2306.07899).

von Oertzen, Christine (2017): "Machineries of Data Power. Manual Versus Mechanical Census Compilation in Nineteenth-Century Europe." In: Osiris 32/1, pp. 129–150.

Wang, Yizhong/Kordi, Yeganeh/Mishra, Swaroop/Liu, Alisa/Smith, Noah A./Khashabi, Daniel/Hajishirzi, Hannaneh (2023). "Self-Instruct: Aligning Language Models with Self-Generated Instructions." In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 13484–13508.

Weiss, Karl/Khoshgoftaar, Taghi M./Wang, DingDing (2016): "A Survey of Transfer Learning." In: Journal of Big Data 3/1: 9.

Xiao, Wen/Xie, Yujia/Carenini, Giuseppe/He, Pengcheng (2023): ChatGPT-steered Editing Instructor for Customization of Abstractive Summarization, arXiv Preprint (https://doi.org/10.48550/arXiv.2305.02483).

Zhang, Daniel/Mishra, Saurabh/Brynjolfsson, Erik/Etchemendy, John/Ganguli, Deep/Sellitto, Michael/Shoham, Yoav/et al. (2021): The AI Index 2021 Annual Report, Stanford, CA: Human-Centered AI Institute, Stanford University (https://arxiv.org/pdf/2103.06312).

Zhao, Shuyin (2023): “GitHub Copilot Now Has a Better AI Model and New Capabilities.”, February 14, 2023 (https://github.blog/2023-02-14-github-copilot-now-has-a-better-ai-model-and-new-capabilities/).

Zheng, Lianmin/Chiang, Wei-Lin/Sheng, Ying/Zhuang, Siyuan/Wu, Zhanghao/Zhuang, Yonghao/Lin, Zi/et al. (2023): Judging LLM-as-a-judge with MT-Bench and Chatbot Arena, arXiv Preprint (https://doi.org/10.48550/arXiv.2306.05685).

Zhou, Viola (2023): “AI is already taking video game illustrators’ jobs in China.”, April 11, 2023 (https://restofworld.org/2023/ai-image-china-video-game-layoffs/).

Zimmerman, Andrew (1999): “Anti-Semitism as Skill: Rudolf Virchow’s “Schulstatistik” and the Racial Composition of Germany.” In: Central European History 32/4, pp. 409–429.