Introduction to the First Workshop on Humanities-Centred Artificial Intelligence

Sylvia Melzer[1,2], Jost Gippert[1,3], Stefan Thiemann[1] and Hagen Peukert[1]

[1]*Universität Hamburg*
[2]*Universität zu Lübeck*
[3]*Goethe-Universität Frankfurt*

Artificial Intelligence (AI) has become a buzz word twirling across the world of science. There is rarely any academic field that would abstain from the progress that AI suggests for each of the scientific disciplines. Given the huge variety of scientific objectives, each discipline has a different set of expectations, preconditions, contributions and views concerning AI and what it should do, what it is, or e.g. where its boundaries are. AI, defined "as the study of agents that receive percepts from the environment and perform actions" [1], does not simply perform routine tasks that are easy to operationalize, but carries out tasks whose execution was traditionally considered to be an exceptional achievement of the human brain.

The contributions to the workshop on Humanities-Centred AI (CHAI) at the 44th German Conference on Artificial Intelligence, are not about solving AI problems but about being able to solve use cases in the field of the Humanities with the use of AI. These papers intend to find a common ground between AI and the Humanities in order to obtain benefits for both fields of research. Though very general, the above perception of AI reveals why it is attractive to an even wider area of potential applications than in the first phase of the digital revolution aiming at the automation of processes that are algorithmically well-understood and easy to specify. While the Humanities remained largely unaffected by last century's digitization, the rapid achievements in AI technology of today can now be harnessed to more profound advances of data analysis in the Humanities. So it is about time to ask: where are possible fields of exploration that are beneficial to research carried out in the Humanities?

The CHAI workshop provided a good entry point to answer this question, by bringing selected topics in recent humanities research together with AI. In the volume at hand, useful and applicable AI approaches for the Humanities are explored. They reach from models of

theoretical computer science, in which agents in clearly defined environments receive stimuli and perform actions over predefined scenarios such as human-aware information retrieval[1] [3] and the intelligent training of systems for non-IT specialists to traditional questions in Linguistics such as word segmentation, addressing concrete assignments to AI and mathematical proves respectively. The contributions aim to reveal that AI is becoming an important part of the Humanities research and thinking alike. AI can be seen as merely a method that makes the life for researchers more convenient by freeing their minds from routine tasks; however, it can also be perceived as a generator of new insights impossible to achieve with the existing inventory of methods. In addition, AI applications can be regarded as both a critic, testing the hypotheses and results of the past, and a catalyst or real inventor of genuine ideas. Whatever the role of AI will be in the decades to come, it will have impact and this should be recognized right from the beginning as part of Humanities' theorizing.

The first contribution addresses a long-standing and still unsolved issue in a linguistic domain, namely word segmentation, and exploits hidden regularities in the input language by means of AI technology. The paper suggests a mechanism to resolve a paradox: how can word boundaries be defined if no previous material is given to learn from – a scenario that all infants encounter and master early in their lives independent of the linguistic environment. Hence, the method is cross-checked with a control language, which could not confirm the universality of the approach. Still, AI has helped to push research in a direction where some aspect of word segmentation, i.e., the role of the input, is better understood and gives eligible hope to bring the field closer to a viable solution.

The second contribution shows how to increase the quality of process models with process mining, a young interdisciplinary research field between machine learning and data mining. The combination of the AI related methods used in the paper has the aim to study discourse corpora by extracting knowledge from raw data, reviewing it, and then using it to improve the processes. The corpus used in the paper contains a collection of telephone conversations between two participants. By extracting the knowledge from the spoken words, various discourse process models are created so that an automatic conversation flow can be derived between specific topics. In the Humanities, this approach can be used, e.g., in evaluating letters to infer which topics are or are not exchanged with which people.

The third contribution presents event and media event extraction as a data mining task that requires natural language understanding. For the extraction tasks several techniques are used, including supervised classification and unsupervised clustering.

The fourth paper argues that AI also anticipates aspects of human information processing and thus requires human knowledge to satisfy a human information need. Humanities scholars who are searching for specific information, e.g., to provide expert testimony on specific artifacts, are obliged to express their information need in the form of one or even more queries to satisfy this information need. The paper shows that the human-aware information seeking approach helps to satisfy information needs by having humans and agents collaborate with each other. In which way this collaboration proceeds is presented in a formal way.

---

[1]The term context-aware in computing was introduced by Schilit, Adams, and Want [2]. Context-aware information retrieval (CAIR) generally describes a search that takes a specific context into account. A context can be document, fairness, risk, location, or simply human.

The fifth contribution gives a new impulse for a new, powerful paradigm for systems that have previously been trained specifically by domain experts, e.g., humanities scholars. The paper elaborates that algorithms for solving the adapted decision or information retrieval problems need to be automatically adapted to support "real" learning while being useful without domain experts having to do this work manually.

These spotlights of ongoing research provide a vivid picture of diverse AI methods and scientific objectives that are useful for the Humanities. They reveal the horizon line of the potential still to come. Even if only little progress seems to be made here, the ideas could point a more reserved mind in the right direction. Flexibility in thinking has always been a unique feature of the Humanities. This does not at all mean the Humanities should subdue themselves to the technological power of AI and its promises, but they should accept it as a natural extension of a flexible mind; as an add-on to the many thinking devices the Humanities have on their workbench. Above all, it would be a great achievement if including AI technologies in the methods inventory of Humanities' disciplines and taking its use as granted without any further ado as part of their natural self-conception were regarded as an impliciteness .

References

[1] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed., Prentice Hall Press, USA, 2009.

[2] B. N. Schilit, N. Adams, R. Want, Context-aware computing applications, in: Proceedings of the Workshop on Mobile Computing Systems and Applications, IEEE Computer Society, 1994, pp. 85–90.

[3] S. Kambhampati, Challenges of Human-Aware AI Systems, arXiv preprint arXiv:1910.07089 (2019).