Challenges in User-Centered Engineering of AI-based Interactive Systems

Jürgen Ziegler

[1] Interactive Systems Group, University of Duisburg-Essen
juergen.ziegler@uni-due.de

Abstract. Intelligent algorithms have reached a new level of performance in recent years and are increasingly employed in application areas such as speech and image recognition, data analytics, or recommender systems. The proliferation of these techniques poses a range of new challenges for the design and engineering of interactive systems since they tend to act as black boxes and do not offer the transparency and level of control to the user which is considered a prerequisite for user-centered design in the HCI field. In this position paper, we provide an overview of the broad areas related to intelligent algorithms and HCI that will need further research in the future to make systems useful, usable and trustable.

Keywords: User Interface Engineering, Interactive Systems, Intelligent Algorithms, Explainable Artificial Intelligence.

1 Introduction

Since the early days of computing, two fundamentally different concepts of the relationship between human and the technical system have been promoted by the fields of Artificial Intelligence (AI) and Human-Computer Interaction, respectively. Artificial Intelligence researchers have been aiming at automating and extensively substituting human mental (and physical) performance while the perspective of Intelligence Augmentation, as represented, among others by Douglas Engelbart, saw technology as a means of enhancing human capabilities which lead to the manifold HCI innovations that we have seen over the past decades.

While AI has in the past often been charged with far-reaching promises, often followed by disillusioning results, recent developments have pointed to a breakthrough in the performance of intelligent technologies. New methods of machine learning, particularly deep learning methods based on neuronal techniques (Goodfellow et al., 2016), and the increasing availability of very large data sets ("big data") with which the systems can be trained have played a significant role in these developments (cf. O'Leary, 2013). In recent years, these techniques have produced amazing results in application areas such as natural language understanding (Otter et al., 2018, Cho et al., 2014), image comprehension (Druzhkov & Kustikova, 2016) He et al., 2015), data analytics (Najafabadi et al., 2015)), recommender systems (Zhang et al., 2019, Donkers, Loepp,

& Ziegler, 2017), or the generation of photorealistic or artistic artifacts (Elgammal, Liu, Elhoseiny & Mazzone, 2017).

While the potential for automating human performance has increased significantly as a result of these successes, the new intelligent techniques also deepen the gap between user and system, since transparency and traceability of decisions generated by probabilistic techniques are mostly not given (Samek, Wiegand & Müller, 2017). Moreover, users have typically no means for interactively controlling the reasoning process. Especially with the otherwise very successful neural network techniques, the system represents a black box from the user's point of view. For system design and development methods that are oriented towards the goals and needs of the users, this represents a considerable challenge, which developers will increasingly have to face up to in the future.

In the following, we will discuss what we see as the main challenges for engineering interactive systems that are fully or partially based on intelligent algorithms. Each challenge stands for a broad range of research questions that have only initially been investigated so far. In general, the challenges ask for a much closer interdisciplinary cooperation between researchers in HCI and AI which up to now are still often separated by a large gap.

2 Challenges for Engineering AI-based Systems

2.1 Presentation and interaction

For the presentation of and interaction with AI components, many specific questions arise beyond the classical aspects of user interface design. These include the choice of a suitable interaction style, ranging from conventional GUI solutions to conversational and language-based interaction to anthropomorphic virtual agents (Lucas et al., 2014). In addition, embedded interactions, as they occur in the Internet of Things, in robotics or in automated vehicles, have to be considered (Kranz, Holleis & Schmidt, 2010). In many cases, the interaction is multimodal, with an extension of previously pursued approaches to multimodality in the context of new machine-learning processes. The increasingly proactive role of AI components, which can initiate situational interactions, poses a variety of research questions, among others with regard to attention guidance, the interruptibility of user activities, or the change of initiative in dialogue.

2.2 Transparency and Explainability

One of the central problems with regard to the usability of intelligent algorithms is the lack of transparency and explanation of automatically made decisions or recommendations (cf. Tintarev & Masthoff, 2015; Donkers et al. 2018). While with conventional techniques such as rule-based systems or decision trees the conclusions can at least in principle still be presented in a comprehensible way, this is usually not the case with neural techniques. In this case, even the developers can hardly comprehend the machine decisions. For reasons of usability as well as information self-determination and data protection (see e. g. European General Data Protection Regulation), it is therefore essential to develop mechanisms that can at least explain the decision mechanisms relevant from the user's point of view and their data basis (Doran, Schulz & Besold, 2017). Although there are initial technical approaches, the question of the type and level of detail of system-generated explanations and their presentation is still a largely open research question (cf. Dosilovic 2018). Overall, the question of user-oriented explanations for machine-made decisions forms a central aspect for the user life and acceptance of AIs.

2.3 User control of intelligent algorithms

In addition to the explanatory nature of algorithmic decisions, the further question from the user's perspective is whether and how algorithmic decision-making processes can be influenced and controlled. In principle, user control can start at different stages of the process. This ranges from the selection of data to be used for learning processes to the selection and parameterization of algorithms and feedback on proposals and decisions made by the system (Amershi et al., 2014). One example is the control of probabilistic models applied in recommender systems through textual tags associated with the items to be recommended (Loepp et al. 2019).

Future research activities on these topics should include the design of interactive methods for the control of algorithmic processes as well as the development of new algorithms that allow more far-reaching intervention possibilities and closer interaction between user and algorithm.

2.4 Human-system cooperation

The old question of the division of tasks between humans and technical systems arises in a new way in the context of AI techniques, as increasingly complex mental tasks can be automated. From the point of view of human-computer interaction, AI systems should support users in their tasks in the best possible way and not pursue a pure substitution strategy. Their use raises fundamental questions regarding the agency and autonomy of users and requires new solutions regarding the interaction and cooperation between user and system. This requires research in areas such as mixed-initiative interactions, handover processes (e.g. automated driving) and mutual learning in collaborative work (e.g. robotics, cf. Lemaignan, 2017).

2.5 Ethical and legal aspects

When machines take on cognitive tasks, a variety of questions arise that would also arise in a comparable form for human actors in the social context (Bostrom & Yudkowski 2011). A central problem here, too, is the inexplicability of AI-based decisions, such as in the granting of loans or in application procedures, which is critical both in ethical and legal terms. The demand for the explanation of algorithmically made decisions is also anchored in the European Data Protection Directive, at least in partial aspects, without the implementation having been sufficiently investigated and clarified so far. Algorithmic decisions can (intentionally or unintentionally) be biased or discriminatory for certain groups of persons (Zemel et al., 2013). Research on how such tendencies can be identified and avoided is still largely in its infancy. However, from both an ethical and a legal point of view, it is still widely open to which actors the responsibility and liability for algorithmic decisions can be attributed. Different disciplines are required to solve these problems, but human-computer interaction can make an important contribution to clarifying these questions and to their implementation in interactive systems.

3 Conclusions

The advent of new approaches such as Deep Learning has immensely enhanced the effectiveness of intelligent algorithms which have begun to pervade a wide range of application areas. A large fraction of these systems are interactive, in that they provide services and functions that are meant to support users in performing their tasks. The black-box nature of most of these techniques, however, prevents users from comprehending and controlling the system which may lead to dissatisfaction and mistrust in the system. So far, there are neither viable nor established methods for engineering AI-based interactive systems. A much closer cooperation between the different fields and an integration of their methods will be needed to overcome the current deficits of intelligent systems from the user's point of view.

References

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine*, *35*(4), 105. https://doi.org/10.1609/aimag.v35i4.2513

Bostrom, N., & Yudkowski, E. (2014). The Ethics of Artificial Intelligence. In W. Ramsey & K. Frankish (Eds.), Cambridge Handbook of Artificial Intelligence (S. 316–334). Cambridge University Press.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734. https://doi.org/10.3115/v1/D14-1179

Donkers, T., Loepp, B., & Ziegler, J. (2017). Sequential User-based Recurrent Neural Network Recommendations. Proceedings of the Eleventh ACM Conference on Recommender Systems, 152–160. https://doi.org/10.1145/3109859.3109877

Donkers, T., Loepp, B., & Ziegler, J. (2018). Explaining Recommendations by Means of User Reviews. Joint Proceedings of the ACM IUI 2018 Workshops co-located with the 23rd ACM Conference on Intelligent User Interfaces (ACM IUI 2018), Tokyo, Japan, March 11, 2018. http://ceur-ws.org/Vol-2068/exss8.pdf

Doran, D., Schulz, S., & Besold, T. R. (2017). What Does Explainable AI Really Mean? A New Conceptualization of Perspectives. *ArXiv:1710.00794 [Cs]*.

Došilović, F. K., Brčić, M., & Hlupić, N. (2018). Explainable artificial intelligence: A survey. 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 0210–0215

Druzhkov, P. N., & Kustikova, V. D. (2016). A survey of deep learning methods and software tools for image classification and object detection. Pattern Recognition and Image Analysis, 26(1), 9–15. https://doi.org/10.1134/S1054661816010065

Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative Adversarial Networks, Generating “Art” by Learning About Styles and Deviating from Style Norms. *ArXiv:1706.07068 [Cs]*.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

Kranz, M., Holleis, P., & Schmidt, A. (2010). Embedded Interaction: Interacting with the Internet of Things. *IEEE Internet Computing*, *14*(2), 46–53. https://doi.org/10.1109/MIC.2009.141

Lemaignan, S., Warnier, M., Sisbot, E. A., Clodic, A., & Alami, R. (2017). Artificial cognition for social human–robot interaction: An implementation. *Artificial Intelligence*, *247*, 45–69. https://doi.org/10.1016/j.artint.2016.07.002

Loepp, B., Donkers, T., Kleemann, T., & Ziegler, J. (2019). Interactive recommending with Tag-Enhanced Matrix Factorization (TagMF). International Journal of Human-Computer Studies, 121, 21–41. https://doi.org/10.1016/j.ijhcs.2018.05.002

Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It’s Only a Computer. Comput. Hum. Behav., 37(C), 94–100. https://doi.org/10.1016/j.chb.2014.04.043

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. Journal of Big Data, 2(1), 1. https://doi.org/10.1186/s40537-014-0007-7

O’Leary, D. E. (2013). Artificial Intelligence and Big Data. *IEEE Intelligent Systems*, *28*(2), 96–99. https://doi.org/10.1109/MIS.2013.39

Otter, D. W., Medina, J. R., & Kalita, J. K. (2018). A Survey of the Usages of Deep Learning in Natural Language Processing. arXiv:1807.10854 [cs]. Abgerufen von http://arxiv.org/abs/1807.10854

Samek, W., Wiegand, T., & Müller, K.-R. (2017). Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *ArXiv:1708.08296 [Cs, Stat]*.

Tintarev, N. & Masthoff, J. (2015) Explaining Recommendations: Design and Evaluation. In: Recommender Systems Handbook, Springer, 353–382

Zemel, R., Wu, Y., Swersky, K., Pitassi, T., & Dwork, C. (2013). Learning Fair Representations. *Proceedings of the 30th International Conference on Machine Learning, PMLR,* 28(3), 325-333.

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep Learning Based Recommender System: A Survey and New Perspectives. ACM Comput. Surv., 52(1), 5:1–5:38. https://doi.org/10.1145/3285029