A Human-centric AI-driven Framework for Exploring Large and Complex Datasets

Maria F. Costabile[1], Giuseppe Desolda[1], Giovanni Dimauro[1], Rosa Lanzilotti[1], Daniele Loiacono[2], Maristella Matera[2] and Massimo Zancanaro[3]

[1] *University of Bari Aldo Moro, via Orabona, 4, Bari, Italy*
[2] *Politecnico di Milano, Via Ponzio, 34/5, Milano, Italy*
[3] *University of Trento, Corso Bettini, 84, Rovereto, Italy*

Abstract

Human-Centered Artificial Intelligence (HCAI) is a new frontier of research at the intersection between HCI and AI. It fosters an innovative vision of human-centred intelligent systems, which are systems that take advantage of computer features, such as powerful algorithms, big data management, advanced sensors and that are useful and usable for people, providing high levels of automation and enabling high levels of human control. This position paper presents our ongoing research aiming to extend the HCAI framework for better supporting designers in creating AI-based systems.

Keywords

Human-centered Artificial Intelligence, design framework, pattern

1. Introduction and motivation

The Human-Computer Interaction (HCI) community acknowledges Artificial Intelligence (AI) as very valuable to create digital technology that can greatly empower people. At the same time, there is still a lack of understanding on how to benefit from AI while preserving the system's reliability, safeness, and trustworthiness for humans [21]. AI is usually conceived with a perspective on autonomy: from autonomous behaviour to autonomous decision-making [17]. In several contexts, a high degree of autonomy might be useful (albeit not without risks). Yet, a different approach might be to exploit AI to support and facilitate human beings by augmenting (and valuing) human cognitive abilities rather than replacing them.

By leveraging the new frontiers of research at the intersection between HCI and AI, we are starting to work on foundational research for a new extension of AI, called Human-Centered Artificial Intelligence (HCAI), by shifting the attention from an algorithm-focused view to a human-centered perspective that requires HCI strategies of design and testing [22]. HCAI perspective ([21], [22]) aims to propose methods to design new interaction paradigms that can amplify, augment, and enhance human performance, in ways that make systems reliable, safe, and trustworthy.

A fundamental element in HCAI is the control by end users. As envisaged by a leading AI researcher [18], a new relationship between humans and machines is needed to design machines that are "not just intelligent but also beneficial to humans". End users' control could be fostered by defining AI models and algorithms that can grant transparency to the system behavior, making it easy for the end users (and all the stakeholders in general) to understand and trust the system's decisions, the so-called Model Explainability [15]. Explainability is generally considered to highlight technical features characterising the performance of AI models. Methods proposed in the literature to open black box models identify

Proceedings of CoPDA2022 - Sixth International Workshop on Cultures of Participation in the Digital Age: AI for Humans or Humans for AI? June 7, 2022, Frascati (RM), Italy
EMAIL: maria.costabile@uniba.it (M. F. Costabile); giuseppe.desolda@uniba.it (G. Desolda); giovanni.dimauro@uniba.it (G. Dimauro); rosa.lanzilotti@uniba.it (R. Lanzilotti); daniele.loiacono@poliba.it (D. Loiacono); maristella.matera@polimi.it (M. Matera); massimo.zancanaro@unitn.it (M. Zancanaro)
ORCID: 0000-0001-8554-0273 (M. F. Costabile); 0000-0001-9894-2116 (G. Desolda); 0000-0002-4120-5876 (G. Dimauro); 0000-0002-2039-8162 (R. Lanzilotti); 0000-0002-5355-0634 (D. Loiacono); 0000-0003-0552-8624 (M. Matera); 0000-0002-1554-5703 (M. Zancanaro)

useful explainability strategies [11]. However, important issues remain open. Different scientific communities address explainability from different perspectives. The explanations provided by the AI community are primarily directed to AI specialists. The HCI community considers these approaches inadequate since they are not meant for the end users, who are possibly domain experts but not AI experts ([14], [23]).

Reliable user control on the system can be also achieved by granting an interactive manipulation of the relevant parameters determining the system behaviour [19]. In the HCAI vision, user control and system autonomy are not considered as opposing each other but rather as two dimensions to be adequately calibrated when designing intelligent systems beneficial to people [21]. Users should be enabled to take advantage of the power of the AI algorithms, but the importance of the knowledge that users, as domain experts, possess must not be neglected. As an example, in [2] an ML-based tool to visually retrieve medical images (tissue from biopsies) from past patients is presented. The tool supports medical decisions with new patients, empowering the physician to cope with the search algorithm on-the-fly, communicating what types of similarities are most important in different situations. This interaction between the human and the system determines a step-wise-refinement that increases the diagnostic utility of images found, as well as the user's trust in the algorithm.

The opportunity for the users to modify the system behavior and adapt it to their needs (the so-called End-User Development [1], [8]), possibly acting on the system's AI models, is crucial in the long term for real empowerment in the use of AI systems. Meta-design principles [8] must be adopted to define a methodological framework in which developers and AI specialists do not design a rigid system, rather they provide a scaffolding environment where adequate model explanations can empower domain experts to reconfigure algorithm outcomes through negotiation. In [12] the authors discuss Interactive ML techniques enabling model reconfiguration through experts' intervention in medical scenarios. However, they also highlight the need for HCI methods to identify adequate interaction paradigms.

Despite the novel contributions mentioned above, a comprehensive framework able to support designers in creating HCAI systems is still lacking. Our goal is to fill this gap by proposing a conceptual framework and a pattern language for the design of a new class of tools with which human beings of different levels of expertise can engage in interactive and iterative meaning-making activities for exploring large and complex datasets. Specifically, the framework promotes i) a new notion of explainability, which can be adequate for non-technical people, and ii) new paradigms for the interaction between humans and systems that can enable a negotiation process that empowers the human not only to understand the reasons that determine a specific system behaviour, but also to modify it through iterative reconfigurations.

2. Three overarching strategies of interaction with AI systems

We aim to extend the HCAI conceptual framework for the design of a new class of tools with which human beings of different levels of expertise can engage in interactive and iterative meaning-making activities for exploring large and complex datasets. Three basic strategies of interaction with AI (ML-based) systems will be explored. They are:

1. *Clarification*: means explaining directly the reason for the behaviour/decision of the AI system. It implies strategies for different types of explanations as an alternative to the so-called "black-box" of deep learning systems (e.g., [20]). The explainability It also involves design patterns to communicate the level of certainty of the system suggestions, which are often based on applying thresholds to computed probabilities (e.g., [3], [16]);
2. *Negotiation*: means reaching the system outcome through a sequence of iterative negotiation steps driven by the interaction between the user and the AI system. It involves meta-strategies and design patterns aimed at segmenting the decision process in progressive steps and providing means to assess and recalibrate the request to the system (e.g., [2]);
3. *Reconfiguration*: means the possibility to trigger simply (within the reach of non-expert users) the retraining of the AI model(s) on new examples. This may happen when the users detect wrong suggestions by the systems or when new data should be incorporated into the model after a negotiation phase (e.g., [12]).

While each one of these strategies has been addressed by some other works, a comprehensive framework that integrates them is still lacking. Also, while several validated patterns exist for standard graphical user interfaces (e.g., [24]), there are still no comparable guidelines for novel paradigms of Human-AI system interaction.
As one of the major outcomes, our research will define interaction-design patterns for HCAI. Patterns are indeed a useful cognitive tool that captures human expertise to provide effective and reusable knowledge [24] on how to manage specific aspects of a design. One of the main features of design patterns is their uniform structure and specification format, which is composed of several items giving the designer indications on the addressed problem, the proposed solution, the context in which the pattern can be applied, the purposes and the benefits expected, and, eventually, examples of the application of the selected pattern. Some of the authors of this position papers have a large experience in the definition of patterns [4], [5], [10], [13].

3. The case study

The three broad strategies above will be experimented within one main case study in the medical domain to derive grounded principles of human interaction with AI systems.

The case study will be in the challenging context of medical cytology, specifically by tackling the problem of AI-supported microscopic analysis of cells contained in the nasal mucosa [6], [7]. We will base our exploration on an existing tool called Rhino-Cyt [6] and we will design, deploy and assess new interface prototypes and related interaction paradigms based on the current ML model and backend system. Unlike what happens for example in the field of haematology, nasal cytology does not yet benefit from a network of public or private laboratories that carry out in-depth analyses quickly and at a low cost. Therefore, the whole diagnostic process is mainly based on direct observation under the microscope that requires prolonged effort by domain experts. The modern scanning systems for cytological preparations and the new affordable digital microscopes enable the design of software systems to support the physician's activity. Rhino-Cyt encodes a deep neural network to automatically identify and classify cells present on a nasal cytological preparation based on a digital image of the preparation itself. Concerning the more standard approach to automating cell counting, Rhino-Cyt aims to move from the current semi-quantitative estimation to a quantitative one, which is more precise and valuable on a scientific level for standardisation, to catalogue cellular elements and get a more accurate diagnosis in the shortest time. These changes may help in the more pervasive use of nasal cytology, a diagnostic investigation that is yet not massively adopted by the new generation of physicians. Although the tool already includes over 3,000 images, the main obstacle at present is the lack of data: some classes are underrepresented and generate underfitting. In this respect, it is necessary to design algorithms that can lead to the achievement of the final diagnosis through a process of successive steps, with the intervention of the physician. In the current interface, the physician manually disambiguates all unclassified or incorrectly classified cells. The physician can also intervene to correct the system's classification if it is regarded as incorrect. Yet, neither specific strategies are implemented at the interface level nor in the general interaction with the system, therefore it is the full responsibility of the physician to decide if and how to intervene.

With the aim of investigating different interaction strategies, we will propose several redesigns of the Rhino-Cyt interface and, in general, of the interaction between the physician and the AI backend along the broad strategies above. Through the assessment of these solutions with the domain experts, we will distil lessons that can help identify patterns for the interaction with AI-enhanced systems.

4. Conclusions

The aim of our current research is to inquire about foundational aspects of Human-Centred AI and its main goal is to deliver a general framework and a library of patterns for interaction with AI tools meant as an extension of the Human-Centred AI framework.

The applied methodology is "research-through design" approach [26], which emphasises the production of prototypes as vehicles for inquiring about foundational aspects of a research challenge. This approach is commonly used in HCI while it is rarely used in AI. Other relevant outcomes of the project are a redesign of the existing system Rhino-Cyt and a set of specific lessons learned for AI in medicine derived from our fieldwork with physicians and specialists in the field of cytology.

5. References

[1] Ardito, Carmelo, Buono, Paolo, Costabile, Maria F., Lanzilotti, Rosa, and Piccinno Antonio. "End users as co-designers of their own tools and products." Journal of Visual Languages & Computing 23.2 (2012): 78-90.

[2] C.J. Cai, E. Reif, N. Hegde, J. Hipp, B. Kim, D. Smilkov, M. Wattenberg, F. Viegas, G.S. Corrado, M.C. Stumpe, M. Terry, Human-Centered Tools for Coping with Imperfect Algorithms During Medical Decision-Making, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), ACM Press, New York, NY, Paper 4, 1–14. doi: 10.1145/3290605.3300234.

[3] Convertino, Gregorio, Zancanaro, Massimo, Piccardi, Tiziano, and Ortega Felipe. "Toward a mixed-initiative QA system: From studying predictors in Stack Exchange to building a mixed-initiative tool." International Journal of Human-Computer Studies, 99 (2017): 1–20.

[4] M.F. Costabile, M. Matera, Guidelines for Hypermedia Usability Inspection, IEEE Multimedia 8.1 (2001) 66-69. doi: 10.1109/93.923955.

[5] F. Daniel, M. Matera, Mashups - Concepts, Models and Architectures. Data-Centric Systems and Applications, Springer, 2014. ISBN 978-3-642-55048-5.

[6] G. Dimauro, F. Girardi, M. Gelardi, V. Bevilacqua, D. Caivano, Rhino-Cyt: A System for Supporting the Rhinologist in the Analysis of Nasal Cytology, in: Proceedings of International Conference on Intelligent Computing (ICIC'18) Springer, 619-630. doi: 10.1007/978-3-319-95933-7_71

[7] Dimauro, Giovanni, Ciprandi, Giorgio, Deperte, Francesca, Girardi, Francesco, Ladisa, Enrico, Latrofa, Sergio, and Gelardi Matteo. "Nasal cytology with deep learning techniques." International Journal of Medical Informatics, 122 (2019): 13–19.

[8] G. Fischer, D. Fogli, A. Piccinno, Revisiting and Broadening the Meta-Design Framework for End-User Development. In Paternò & Wulf (Eds) New Perspectives in End-User Development (2017). Springer. doi:10.1007/978-3-319-60291-2_4

[9] E. Giacomello, P.L. Lanzi, D. Loiacono, L. Nassano, Image Embedding and Model Ensembling for Automated Chest X-Ray Interpretation, in: Proceedings of International Joint Conference on Neural Networks (IJCNN '21) IEEE, 1-8. doi: 10.1109/IJCNN52387.2021.9534378

[10] M. Gianotti, F. Riccardi, G. Cosentino, F. Garzotto, M. Matera, Modeling Interactive Smart Spaces. In Conceptual Modeling: 39th International Conference (ER 2020). Springer-Verlag, Berlin, Heidelberg, 403–417. doi: 10.1007/978-3-030-62522-1_30.

[11] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A Survey of Methods for Explaining Black Box Models, ACM Comput. Surv. 51, 5 (2018) Article 93, 42 pages. doi: 10.1145/3236009.

[12] Holzinger, Andrea. "Interactive machine learning for health informatics: When do we need the human-in-the-loop?" Brain Informatics, 3.2 (2016): 119–131.

[13] Lanzilotti, Rosa, De Angeli, Antonella, Ardito, Carmelo, and Costabile, Maria F. "Do patterns help novice evaluators? A comparative study." International Journal of Human-Computer Studies, 69 (2011): 52-69.

[14] Q.V. Liao, D. Gruen, S. Miller, Questioning the AI: Informing Design Practices for Explainable AI User Experiences, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). ACM, New York, NY, 1–15. doi:10.1145/3313831.3376590.

[15] S.M. Lundberg, S. Lee, A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 4768–4777.

[16] Z. Lu, M. Yin. Human Reliance on Machine Learning Models When Performance Feedback is Limited: Heuristics and Risks, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). ACM, New York, NY, Article 78, 1–16. doi: 10.1145/3411764.3445562.

[17] F. Paternò, M. Burnett, G. Fischer, M. Matera, B. Myers, A. Schmidt, Artificial Intelligence versus End-User Development: A Panel on What Are the Tradeoffs in Daily Automations?, in: Proceedings of INTERACT 2021, Springer.
[18] Russel S. Human Compatible - Artificial Intelligence and the Problem of Control. Viking, 2019.
[19] A. Schmidt, Interactive Human Centered Artificial Intelligence: A Definition and Research Challenges, in: Proceedings of the International Conference on Advanced Visual Interfaces (AVI 20) ACM, New York, NY, USA, Article 3, 1–4. doi: https:10.1145/3399715.3400873.
[20] M. Setzu, R. Guidotti, A. Monreale, F. Turini, D. Pedreschi, F. Giannotti, Glocalx-from local to global explanations of black box AI models. Artificial Intelligence (2021) 294. doi: 10.1016/j.artint.2021.103457
[21] B. Shneiderman. "Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. Int. J. Hum. Comput. Interact. 36.6 (2020): 495-504.
[22] B. Shneiderman, Human-Centred AI. Oxford University Press, 2022.
[23] H. Suresh, S.R. Gomez, K.K. Nam, A. Satyanarayan, Beyond Expertise and Roles: A Framework to Characterize the Stakeholders of Interpretable Machine Learning and their Needs, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). ACM, New York, NY, Article 74, 1–16. doi: 10.1145/3411764.3445088.
[24] J. Tidwell, C. Brewer, A. Valencia, Designing Interfaces Patterns for Effective Interaction Design. O'Reilly, 2020.
[25] J. Zimmerman, J. Forlizzi, S. Evenson, Research through design as a method for interaction design research in HCI, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07). ACM, New York, NY, 493–502. doi: 10.1145/1240624.1240704.