HHAI 2024: Hybrid Human AI Systems for the Social Good
F. Lorig et al. (Eds.)

doi:10.3233/FAIA240229

Operational Criteria of Hybrid Intelligence for Generative AI Virtual Assistants

Jacob SHERSON[1], Janet RAFNER and Safinaz BÜYÜKGÜZEL
Center for Hybrid Intelligence, Dep. of Management, Aarhus University
ORCiD ID: Jacob Sherson https://orcid.org/0000-0001-6048-587X, Janet Rafner https://orcid.org/0000-0001-9264-3334, Safinaz Büyükgüzel https://orcid.org/0000-0002-4692-5143

Abstract. The concept of Hybrid Intelligence (HI) is frequently used interchangeably with Human-Centered AI (HCAI) and more broadly as human-in-the-loop. Dellerman et al. [1] outlined three differentiation criteria, emphasizing in particular the need for an evolving continuum of human-AI learning, a concept that has proven challenging to operationalize effectively. Recent efforts aim to expand the definition of HI beyond the domain of human-computer interaction to include application-oriented insights from management science [2]. This broader perspective integrates vital components such as facilitating end-user co-creation through narrative frameworks that foster psychological safety by addressing fears of job displacement [3,4], mitigating risks of deskilling during system deployment and scaling [5], and supporting business process innovation [2]. Additionally, in contrast to HCAI, the name hybrid intelligence conveys the possibly symmetric human-machine relationship and thereby preserves some of the disruptive potential of automated AI rather than relying on purely augmentation of human tasks and intentions [3]. Explicitly, the HI interaction should not only augment the existing, predefined task but also support aspects such as (business) process and business model re-engineering. Despite these considerations, a thorough discussion on which of the many established HCAI concepts and design guidelines form crucial components in achieving the aims of HI has so far been absent in literature. In particular, as it is becoming more and more likely that most knowledge workers will within a short timeframe become operators of complex virtual assistants tapping into LLMs and natural language interfaces, it becomes urgent to ensure that the human-ai interface and associated narrative is constructed to support HI principles and objectives. To initiate this discussion, we formulate explicitly updated HI design criteria in particular for generative AI virtual assistant design and discuss relevant HCAI concept.

Keywords. Keyword, keyword

1. HI Design Criteria

1: Interaction objective: Design an interaction that empowers users to reimagine and **innovate workflows** as well as entire **task purposes** and intended value generation streams while preventing long term effects of employee **deskilling** due to excessive technological reliance.

[1] Corresponding Author: Jacob Sherson, sherson@mgmt.au.dk.

2: Interaction narrative: Incorporate these human-centered objectives explicitly into the overall **application narrative** and visual design to facilitate **psychological safety** ensuring **end-user** engagement into the **co-creation** of the design process as well as adaptation and continuous innovation after tool deployment. Note, a use case is only truly hybrid intelligent if it can be used as a case for a **positive Future of Work for humanity**.
3: Interaction design: Design for a continually evolving **mutual human-AI learning** while realizing high degrees of **automation and sense of human control.**

The aim is to evolve HI design into a prescriptive set of design principles that application developers can strive to achieve as well as allow for concrete discussions of phenomena that are absent in order to differentiate clearly from the more generically applied concepts such as human-in-the-loop.

Some **design concepts and challenges** that need to be deliberately and contextually designed for in a hybrid intelligence VAs include:

Narrative explicitation and Interactive visualization to incorporate human concerns into an HCAI loop [6]. This includes taking concerns such as fear of job loss explicitly into the naming and design of virtual assistants. For instance, although presented with a friendly graphic the name Einstein GPT of the Salesforce virtual assistant may not instill a sense of equal partnership: when a query fails, whose fault is it? Einstens or the potentially technologically inexperienced and fearful end user?

Seamless vs seamful design: whereas seamless interaction can enhance a sense of user-flow, it can also hide algorithmic complexities within a black box. Conversely, within seamful design one can anticipate potential breakdowns and exploitatively craft custom solutions and procedural shortcuts connecting and traversing seams [7]. An application of this would be transforming the "undo-button event" in which the user cancels an attempt of automation from a frustrating interruption into a positive, collective learning experience.

Navigate the **transition from producer to evaluator** to avoid productivity loss by i.a. (i) continuous feedback, (ii) system personalization, (iii) ecological interface design, (iv) main task stabilization and timing, and (v) clear task allocation [8]. A critical HI design challenge involves facilitating users to navigate and experiment with task flows and loops that were unfeasible before Generative AI due to the extensive manual effort required. For example, shifting the focus from detailed writing processes to more strategic planning and outlining activities.

Convergent vs divergent support: Designers explicitly need to consider on a case-by-case basis if the assistant should support divergent or convergent thinking by presenting options or suggested solutions. This could involve showing several options for the next step vs showing/completing several steps along the workflow. Here, attention should in particular be given to the pleasing vs provoking tradeoff either fulfilling expectations or instilling new ideas [9]. The user provocation can be particularly relevant for overcoming confirmation bias and, in the HI context, in supporting task objective re-engineering.

Level of foresight: The phenomenon in human-AI interaction where the AI anticipates the user's desires based on ambiguous input and performs several automated steps in one action is referred to as proactive automation or predictive interaction [10]. This involves the AI system analyzing patterns in user behavior, preferences, and context to predict and execute tasks without explicit instructions at every step, aiming to streamline the user experience by reducing the need for direct commands or inputs.

Acknowledgements: The research project was supported by Copenhagen Fintech as part of the "National Position of Strength programme for Finans & Fintech" funded by the Danish Ministry of Higher Education and Science" and Aarhus University Research Foundation.

References

[1] Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. Business & Information Systems Engineering, 61(5), 637-643.

[2] Rafner, J., Bantle, C., Dellermann, D., Söllner, M., Zaggl, M. A., & Sherson, J. (2022a). Towards Hybrid Intelligence Workflows: Integrating Interface Design and Scalable Deployment. In HHAI2022: Augmenting Human Intellect (pp.310-313). IOS Press.

[3] Holgersen, L., Rogozinska, P., Dellermann, D., Sherson, J.F. (2023) AI or Hybrid Intelligent Future of Work? Controlled experiment on narrative importance, in submission, https://urlis.net/5xuou39r

[4] Mao, Yaoli, et al. A hybrid intelligence approach to training generative design assistants: partnership between human experts and AI enhanced co-creative tools. HHAI 2023: Augmenting Human Intellect. IOS Press, 2023. 108-123.https://doi.org/10.3233/FAIA230078

[5] Rafner, J., Bantle, C., Dellermann, D., Söllner, M., Zaggl, M. A., & Sherson, J. (2022a). Towards Hybrid Intelligence Workflows: Integrating Interface Design and Scalable Deployment. In HHAI2022: Augmenting Human Intellect (pp.310-313). IOS Press.

[6] Hoque, M. N., Shin, S., & Elmqvist, N. (2024). Harder, better, faster, stronger: Interactive visualization for human-centered AI tools. arXiv. https://arxiv.org/pdf/2404.02147.pdf

[7] Ehsan, U., Liao, Q. V., Passi, S., Riedl, M. O., & Daume III, H. (2022). Seamful XAI: Operationalizing Seamful Design in Explainable AI. arXiv preprint arXiv:2211.06753.

[8] Simkute, A., Tankelevitch, L., Kewenig, V., Scott, A. E., Sellen, A., & Rintel, S. (2024). Ironies of Generative AI: Understanding and mitigating productivity loss in human-AI interactions. arXiv preprint arXiv:2402.11364.

[9] Rezwana, J., & Maher, M. L. (2023). Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems. ACM Transactions on Computer-Human Interaction, 30(5), 1-28.

[10] Heer, J., Hellerstein, J. M., & Kandel, S. (2015). Predictive interaction for data transformation. In Proceedings of the 7th Biennial Conference on Innovative Data Systems Research (CIDR '15). Asilomar, California, USA. http://creativecommons.org/licenses/by/3.0/