Human Interaction and Emerging Technologies (IHIET-AI), Vol. 201, 2026, 294–303
https://doi.org/10.54941/ahfe1007183

Augmented Cognition Requires a Psychologically Sound Human Role: A Methodical Approach

Patrick Zinsli, Stephanie Kalt, Nerissa Dettling, Samira Hamouche, and Toni Waefler

University of Applied Sciences and Arts Northwestern Switzerland (FHNW), School of Applied Psychology, 4600 Olten, Switzerland

ABSTRACT

This paper presents a method for optimizing the allocation of functions in human-AI systems. The main objective is to provide humans with a psychologically meaningful role that avoids the negative effects of automation, such as complacency or deskilling. To this end, the method comprises three steps: (i) actual requirements are analysed through a thorough task analysis, (ii) potential AI functions and their impact on the role of humans are described and assessed using human-centred evaluation criteria, and only after a psychologically meaningful human role has been achieved (iii) is the interface for collaboration between humans and AI designed. This is because even good interface design cannot correct situations in which people are assigned a role that does not meet critical psychological requirements and is therefore impossible for them to perform. The paper presents the method using a fictitious example.

Keywords: Human-AI system, Human-AI function allocation, Task analysis, Human role design, Human-AI collaboration design

INTRODUCTION

The capabilities of AI and the quality of AI-generated output are increasing at an unprecedented rate. At the same time, the challenges of human-AI collaboration are also growing. This is because the better an AI performs, the more difficult it becomes for humans to recognise AI malfunctions. For example, while hallucinations from a poor LLM are quite obvious and therefore easy to identify, hallucinations from a powerful LLM are sophisticated and opaque, making them increasingly difficult for humans to detect. Challenges in human-AI collaboration as described e.g. by Endsley (2023) are therefore not symptoms of a new technology's teething problems, but rather inherent in AI itself. Essentially, the challenge lies in the fact that humans are expected to act as a firewall for AI deficiencies. However, to supervise an AI that processes much more data humans are capable of by a model humans do not understand, is a task that exceeds human capabilities. As a consequence, humans are not suited to take on the task of supervising AI or evaluating AI-generated recommendations and bearing responsibility for them.

Received October 7, 2025; Revised November 28, 2025; Accepted December 8, 2025; Available online April 1, 2026

Against this background, the HORIZON project AI4REALNET (cf. ai4realnet.eu) aims to research AI-based solutions addressing critical systems (electricity, railway and air traffic management) that are traditionally operated by humans, and where AI systems complement and augment human abilities. As part of the project, the "Supportive AI Framework" (Waefler et al., 2025) was developed. This framework aims at an intensified human-AI collaboration (Waefler, 2021), in which humans are active participants rather than passive observers of AI or recipients of AI-generated information. Rather, humans and AI are considered a joint cognitive system (Hollnagel & Woods, 2005) based on their qualitatively different but complementary strengths and weaknesses. With the aim to augment human cognitive abilities, the framework conceptualises ways for AI to explicitly support human cognitive processes such as decision-making or learning.

This paper covers the methodological part of the "Supportive AI Framework". A method is presented, together with suitable instruments, that supports the analysis and design of human-AI collaboration based on cognitive task analysis. Special attention is paid to the creation of a psychologically coherent human role in human-AI collaboration. This is to avoid the negative consequences of AI as described e.g. by Endsley (2023) or Buçinca et al. (2024) (e.g. deskilling, demotivation or cognitive overstraining). The method provides guidance on creating detailed descriptions of the human's role in specific scenarios, as well as how human and AI collaborate. It also includes a detailed analysis of the (tacit) knowledge and skills that humans need to fulfil their assigned roles, as well as how these are acquired. Both are critical to avoid deskilling.

In the following sections, the method is described in detail and illustrated by fictitious examples where augmented human cognition in knowledge-intensive tasks is envisioned. The aim is to combine humans and AI in the tradition of sociotechnical system design (e.g. Ulich, 2011) and complementary function allocation (e.g. Waefler et al., 2003).

DESIGNING FOR HYBRID INTELLIGENCE

The method consists of three interrelated steps that systematically guide the development of effective human-AI interaction:

- Step 1, as-is analysis: Descriptive analysis of the task to be supported by AI.
- Step 2, role design: Description and normative evaluation of the roles assigned to humans and AI.
- Step 3, collaboration design: Description and normative evaluation of human-AI collaboration.

These steps are conceptually sequential. However, the process remains iterative, so that findings from later steps can be incorporated into earlier steps. Across all steps of the method, we recommend involving a project team consisting of domain experts, human factors specialists, and AI experts. While domain experts contribute practical insights into real work practices, human factors specialists provide psychological insights, lead data collection and facilitate the workshops. AI experts are integrated from the outset to inform the project team about technical opportunities and feasibility, enabling coherent design decisions in later steps.

The first step is to gain a deep understanding of the task into which the AI is to be integrated. The work analysis aims to develop a descriptive representation of the task within the sociotechnical system in its current state. It is essential to understand how the task is actually performed in practice rather than how it is described on paper. Therefore, the method recommends qualitative research techniques such as observations and interviews with experienced domain experts. Once the task is thoroughly understood and described, the process continues with the design phase.

The design phase comprises two steps: role design and collaboration design. Clearly defining the roles of humans and AI is a crucial prerequisite for developing effective human-AI collaboration. From a human factors perspective, a well-balanced role distribution ensures that the sociotechnical system leverages human strengths while compensating for human limitations. This step addresses the guiding question: What roles will humans and AI take on in future cooperation?

Once the roles have been clearly defined, the next step focuses on designing the concrete collaboration of humans and AI. This step addresses the question: How do humans and AI collaborate? The goal is to design the collaboration in such a way that humans receive all relevant information from the AI in an appropriate form, enabling informed decision-making and effective task execution.

Both design steps comprise a descriptive part and a normative part. This distinction is crucial, as it provides an objective and standardised procedure that transforms subjective impressions into measurable factors, ensuring a consistent and normative basis for further development.

For the normative part, the method is designed to allow the application of different evaluation criteria. For step 2 – role design – we specifically developed evaluation criteria for “knowledge” and “learning” (see below, section Role Design). However, we recommend applying further evaluation criteria for role design (e.g. Clegg, 2000; Waefler et al., 2003) as well as for collaboration design (e.g. Amershi et al., 2019). Regardless of which criteria are chosen, it is essential to collect corresponding information during the as-is analysis.

The following section outlines the three steps of analysis and design in more detail.

Step 1: As-Is Analysis

Once the task under analysis has been defined, its boundaries – where it begins and ends – must be clearly defined in order to determine the scope of the analysis and identify the relevant participants. A wide range of data collection methods can be applied, including interviews, focus groups, and observation-based techniques. In our method, obtaining a detailed description of the individual activities within knowledge-intensive tasks is particularly relevant, including all subtasks, necessary to achieve the overarching task objective. Therefore, four main guiding questions need to be addressed:

1. What is the objective of the task and what subtasks are required to achieve this objective?
2. How are the task and its subtasks carried out in practice?
3. How is the task embedded into the organization, i.e. what kind of collaboration takes place within the organisation?
4. What data is required to apply the selected evaluation criteria?

The results of the as-is analysis are summarised in a table. It provides a comprehensive overview of the various subtasks necessary to complete the overall task. However, the list of subtasks does not necessarily reflect a chronological sequence. It is also possible that certain subtasks within the task are performed repeatedly, but are still only described once in the table. For each subtask, a brief description should be provided, including its objective and organisational context. Furthermore, all relevant information for applying the selected evaluation criteria should be included into the table.

Table 1: An example of an "as-is table" from an HVAC company for the task "preparing quotations." The columns present two illustrative subtasks as examples, while the rows represent the descriptive categories from step 1 as well as the information required for applying the evaluation criteria "knowledge" and "learning."

Subtask	Analyse Project Data	Planning HVAC-Systems
Description	Review project documentation and relevant data from previous projects to define system requirements.	Develop and select an HVAC system using insights derived from analysed project data.
Objective	Developing a knowledge base for informed planning decision to define HVAC system requirements.	Create a technically feasible HVAC system meeting project and operational requirements.
Human task	Collect and synthesise building data (e.g., size, floor plan, location, energy consumption) and data from previous projects; analyse the data to identify patterns, potential risks, and best practices; evaluate technical, regulatory, and legal factors affecting system design; define system requirements	Develop possible HVAC systems; evaluate the alternatives, select the optimal system; verify system feasibility through calculations and expert review
Knowledge	Contextual knowledge: project scope, constraints, and historical data Interpretation of the collected information (e. g. pattern recognition) and implications for planning.	Experience-based knowledge of HVAC behaviour, operational challenges, planning errors, performance, reliability, and lessons from previous system implementations.
Learning	Learning from the analysis of historical project data and feedback from colleagues or customers.	Learning through regular project data analysis, long-term system performance observation, peer review of feasibility, and hands-on installation practice.
Organization	Provide information	Colleagues review technical feasibility

Table 1 shows two subtasks from a fictitious example of an “as-is task analysis” for a company specialised in energy-efficient heating, ventilation and air conditioning (HVAC) systems. It describes the subtasks as they are currently performed, the knowledge required to perform the subtasks, the sources for learning this knowledge, and the organizational context of the subtasks.

The fictional company is experiencing an efficiency problem: preparing quotations requires considerable cognitive and time resources, while existing knowledge from previous quotations remains unused due to its sheer volume. The company aims at improving the quality and accuracy of its quotations. Preparing quotations requires extensive specialist knowledge and is technically demanding. The aim of using AI is to support the quotation process, improve information utilisation, and enhance the quality of related decisions.

The method applied to elaborate this tabular presentation is based on the tradition of Sociotechnical System Analysis (STS) and Hierarchical Task Analysis (HTA). STS emphasises a focus on human work activities, providing insight into working conditions, task design, and their impact on overall system performance (Ulich, 2011). HTA is widely used in the field of human factors and involves describing tasks in a hierarchy of objectives, subobjectives, operations, and plans. The hierarchical structure helps to make the task and its internal relationships visible (Stanton et al., 2013). If needed, additional analyses can be integrated, such as Activity-Centred Task Analysis (ACTA) or Goals, Operators, Methods, and Selection Rules Analysis (GOMS) (cf. Stanton et al., 2013 for an overview of corresponding methods).

In summary, step 1 – as-is analysis – provides a systematic understanding of the current task within its sociotechnical context. By analysing how work is actually performed this step establishes a solid foundation for the subsequent design steps. Building on these insights, the following section focuses on the procedure of step 2: role design.

Step 2: Role Design

The objective of step 2 is to design a human role that is psychologically realistic and meaningful. An iterative procedure serves this purpose. First, for each subtask potential AI functions are identified. Second, the changes that will result for the role of humans due to the implementation of these AI functions in work processes are described. Third, the resulting human roles are assessed using evaluation criteria based on psychology. This process is carried out iteratively until the role of humans, as changed by the implementation of AI, is psychologically acceptable.

The method employed in step 2 refers to user story maps. As a process-oriented visualisation, user story maps present the system requirements in a clear and structured manner. Their broad perspective supports collaboration and helps align the understanding of all stakeholders (Steimle & Wallach, 2023).

The user story map elaborated in step 2 is based on the as-is table created in step 1, with the addition of two rows for potential AI functions. These two rows make a distinction between AI functions supporting humans in task execution versus AI functions automating the task completely and hence replacing the humans. The purpose of this distinction is to encourage participants to consider not only task automation but also explicitly consider possible AI functions that specifically support humans in perfoming their tasks. Subsequently, for each of the possible AI functions it is described, how it changes the human's task. At this point, the user story map evolves from its "as-is" state to a future "to-be" version: the "User Task" row specifies the role of the human, while the "AI Function" row defines the corresponding functions of the AI.

Once an initial version of possible AI functions has been developed, their impact on the human role is evaluated based on selected evaluation criteria. In line with the earlier description of the iterative procedure the specification of the AI functions and the evaluation of the resulting human role are repeated until a psychologically sound human role has been developed. As mentioned above, any evaluation criteria may be applied. However, in this project we developed criteria specifically for motivation (Hamouche et al., 2025) as well as for knowledge and learning. The latter are described in more detail below.

Knowledge: In order to be able to contribute to the fulfilment of the task, humans need corresponding, domain-specific knowledge. However, if AI functions take over parts of a task and humans no longer perform those parts, there is a risk of gradual loss of skills and expertise. Moreover, additional knowledge might be required to be able to critically assess and validate AI generated recommendations. Therefore, the knowledge-related evaluation criterion we propose aims to ensure that the necessary explicit and tacit knowledge continues to be used in daily practice so that it is not lost.

For practical application, we operationalised evaluation anchors divided into three levels – low, medium, and high. Knowledge refers both to the expertise that domain experts need to perform their work and to the preservation of this expertise. The anchors are presented in Table 2.

Table 2: Qualitative evaluation anchors for the criterion "knowledge".

Low	Medium	High
The knowledge relevant to the execution of the subtask is rarely applied and therefore seldom reinforced.	The knowledge relevant to the execution of the subtask is applied occasionally but repeated irregularly.	The knowledge relevant to the execution of the subtask is applied continuously and reinforced regularly.

Table 3: Qualitative evaluation anchors for the criterion "learning".

Low	Medium	High
The person carrying out the task gains minimal concrete experience and is therefore unable to progress through a full cyclical learning process.	The person carrying out the task gains concrete experience, which can inform adaptations of existing ideas or the generation of new ones; however, these ideas are rarely tested in practice.	The person carrying out the task gains concrete experience, reflects on it, and derives potential improvements. The knowledge acquired can be actively tested in practice.

Learning: Learning describes the process through which new knowledge is created, or existing knowledge is further developed. If AI takes part of the task, learning opportunities may be lost. While humans who worked in the field before parts of the task were automated by AI trained their skills and thus developed the relevant expertise, this opportunity could be lost when AI is used. Therefore, the learning-related evaluation criterion we propose aims to maintain opportunities for experiential learning (Kolb, 1984) of critical knowledge. Corresponding evaluation anchors are presented in Table 3.

Table 4: An example of a possible human-AI role design. The columns show the two illustrative subtasks (see also Table 1), while the rows show possible AI functions (supportive and automated) as well as the resulting human task.

Subtask	**Analyse Project Data**		**Planning HVAC-Systems**			
Human task	Instruct the AI to collect and analyse relevant project information	Review and validate the AI-derived system requirements	Develop possible HVAC system options	Formulate hypotheses the evaluate options regarding performance or suitability for each option	Evaluate the AI-provided pros and cons and select the preferred option.	Verify technical feasibility through calculations and expert review.
AI functions (supportive)				Generate pros and cons for each HVAC system option.		
AI functions (automate)	Collects and processes relevant data; identify patterns, risks, and best practices	Generate system requirements based on data analysis.				

Table 4 shows the exemplary subtasks from step 1, as-is analysis (cf. Table 1). The two distinct AI concepts (supportive and automated) and their impact on human tasks are added to the table. In the fictitious example, the automated AI functions for the subtask "Analyse project data" poses challenges concerning the evaluation criterion for knowledge. If the AI fully automates the analysis of building information and generate system requirements, users may no longer acquire the necessary contextual understanding to verify whether the solutions generated by the AI are appropriate. Furthermore, humans can no longer sufficiently learn the knowledge required to perform the subtask "Planning HVAC systems".

In contrast, the AI function for "planning HVAC systems" (see Table 4) is an unproblematic example regarding the evaluation criteria referring to knowledge and learning. According to the concept of Evaluative AI (Miller, 2023), the human formulates hypotheses for each planned system, while the

AI assesses corresponding advantages and disadvantages. This allocation of functions ensures that humans actively use and thus maintain their knowledge. In addition, the evidence provided by AI for or against human hypotheses offers learning opportunities. Further inspiration for potential AI support is provided by the Supportive AI Framework (Waefler et al., 2025).

Once a psychologically sound function allocation is established the as-is table from step 1 is transferred into a to-be user story map. It outlines the requirements for the future AI implementation and serves as the foundation for the next step: developing the collaboration between humans and AI.

Step 3: Collaboration Design

While AI functions and a psychologically sound human role have been developed in the previous step, this step aims at designing human-AI collaboration. To do so, the method applies a user story board to establish a shared understanding among the involved stakeholders and to identify the types of interactions required for effective collaboration.

The user story board represents each subtask in a frame, illustrating the collaboration between humans and AI based on their assigned functions. Additionally, the developed role design is described beneath each frame. Figure 1 presents a user story board for the previous example. It is based on the role design described in step 2.

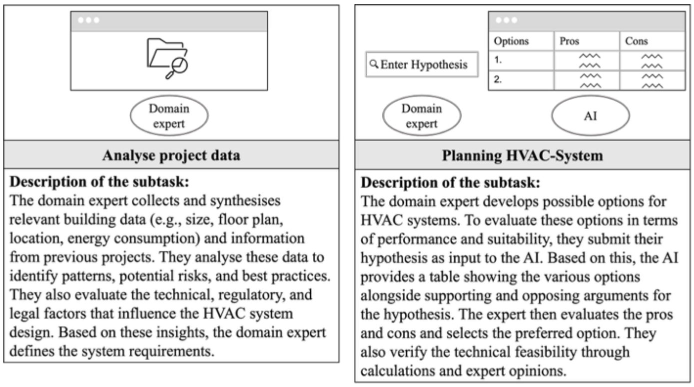

Figure 1: User story board depicting two exemplary subtasks. Each frame illustrates the collaboration design between human and AI based on the role design from Step 2. For subtask "Analyse project data", no AI support is provided. For subtask "Planning HVAC system", the human formulates hypotheses in a text field, while the AI provides evidence for and against these hypotheses in a tabular form.

The initial to-be user story board is iteratively refined in workshops within the project team. To support this process, design templates such as dropdown menus, input fields, and other interface elements should be available during the workshop.

The resulting collaboration design is subsequently evaluated using selected evaluation criteria. Various frameworks exist for assessing human-AI collaboration (e.g. Amershi et al., 2019; Hoffman et al., 2018; Huchler et al., 2020; Schmidt & Herrmann, 2017; Shneiderman et al., 2016).

As in step 2 – role design – the process follows an iterative cycle: an initial solution is developed, assessed against the criteria, and refined accordingly. Multiple iterations should be conducted until the collaboration design reaches a satisfactory level. All artefacts developed throughout the method serve as essential foundations for the subsequent AI development.

CONCLUSION

Based on the traditions of sociotechnical system design (e.g., Clegg, 2000) and human work design (e.g., Ulich, 2011) this paper presents a method for optimising function allocation in human-AI systems. The method takes a complementary design approach, i.e. humans and AI are viewed as qualitatively different with different strengths and weaknesses and are therefore suited to complement each other. A particular focus is placed on designing a psychologically sound human role that promotes their expertise and thus their agency. To do so, potential AI functions are derived from real-world requirements (analysed in step 1) and assessed for their impact on the human's role based on human-centred evaluation criteria (step 2). Only when a thoroughly balanced human-AI function allocation has been achieved, the interface for human-AI collaboration is designed (step 3). The method is being further developed and evaluated as part of the AI4REALNET project.

ACKNOWLEDGMENT

AI4REALNET has received funding from European Union's Horizon Europe Research and Innovation Programme under the Grant Agreement No 101119527 and from the Swiss State Secretariat for Education, Research and Innovation (SERI). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union and SERI. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for Human-AI Interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–13. https://doi.org/10.1145/3290605.3300233

Buçinca, Z., Swaroop, S., Paluch, A.E., Doshi-Velez, F. & Gajos, K.Z (2024). Contrastive Explanations That Anticipate Human Misconceptions Can Improve Human Decision-Making Skills. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2024, Woodstock, NY. ACM, New York, NY, USA, 35 pages

Clegg, C. W. (2000). Sociotechnical principles for system design. Applied Ergonomics, 31(5), 463–477. https://doi.org/10.1016/S0003-6870(00)00009-0

Endsley M. R. (2023). Ironies of artificial intelligence. Ergonomics, 66(11), 1656–1668. https://doi.org/10.1080/00140139.2023.2243404

Hamouche, S., Dettling, N., & Waefler, T. (2025). Applying Job Design Criteria for Effective Human-AI Collaboration. Human Interaction and Emerging Technologies (IHIET 2025). https://doi.org/10.54941/ahfe1006695

Hoffman, R.R., Mueller, S.T., Klein, G., and Litman, J. (2018). Metrics for Explainable AI: Challenges and Prospects. Technical Report, DARPA Explainable AI Program.

Huchler, N., Adolph, L., André, E., Bauer, W., Bender, N., Müller, N., ... & Suchy, O. (2020). Kriterien für die Mensch-Maschine-Interaktion bei KI. Ansätze für die menschengerechte Gestaltung in der Arbeitswelt. Plattform Lernende Systeme, München.

Hollnagel, E. & Woods, DA. (2005) Joint cognitive systems: foundations of cognitive systems engineering. CRC Press, Boca Raton.

Kolb, D. A. (1984). Experiential Learning: Experience as the Source of Learning and Development. Prentice-Hall.

Miller, T. (2023). Explainable AI is Dead, Long Live Explainable AI! Hypothesis-driven Decision Support using Evaluative AI. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, 333–342. https://doi.org/10.1145/3593013.3594001

Schmidt, A. & Herrmann, T. (2017). Intervention User Interfaces: Interactions, 24(5), 40–45. https://doi.org/10.1145/3121357

Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., and Elmqvist, N., Designing the User Interface: Strategies for Effective Human-Computer Interaction: Sixth Edition, Pearson (May 2016) http://www.cs.umd.edu/hcil/DTUI6

Stanton, N., Salmon, P., Walker, G., Baber, C., & Jenkins, D. (2013). Human Factors Methods: A Practical Guide for Engineering and Design. In Human Factors Methods: A Practical Guide for Engineering and Design, 2nd Edition (2nd edition). Ashgate Publishing Limited.

Steimle, T., & Wallach, D. (2023). Collaborative UX Design: Lean UX und Design Thinking: teambasierte Entwicklung menschzentrierter Produkte (2., aktualisierte und erweiterte Auflage). dpunkt.verlag.

Ulich, E. (2011). Arbeitspsychologie (7. Auflage). vdf Hochschulverlag AG.

Waefler, T., Grote, G., Windischer, A. & Ryser, C. (2003). KOMPASS: A Method for Complementary System Design. In: Handbook of Cognitive Task Design. E. Hollnagel (Ed.). Mahwah, NJ: Lawrence Erlbaum, pp. 477–502.

Waefler, T. (2021). Progressive Intensity of Human-Technology Teaming. Proceedings of the 5th International Virtual Conference on Human Interaction and Emerging Technologies, IHIET 2021, August 27–29, 2021, France, pp. 28–36.

Waefler, T., Hamouche, S., Eisenegger, A. (2025). The Supportive AI Framework: From Recommending to Supporting. In: Schmorrow, D.D., Fidopiastis, C.M. (eds) Augmented Cognition. HCII 2025. Lecture Notes in Computer Science, vol 15778. Springer, Cham. https://doi.org/10.1007/978-3-031-93724-8_22