



# Psychological Aspects of Human-Machine Interaction

Applied Machine Learning Days (AMLD), EPFL Lausanne

Toni Waefer; February 11th, 2025



AI4REALNET has received funding from [European Union's Horizon Europe Research and Innovation programme](#) under the Grant Agreement No 101119527



[ai4realnet.eu](https://ai4realnet.eu)



# Institute Humans in Complex Systems (MikS)

- Analysis, evaluation and design of complex systems
  - Increased technical and organizational complexity
  - Individuals and groups confronted with such complexity
  
- Objectives
  - Increase the reliability and safety of sociotechnical systems
  - Healthy humans and organizations



# Human Agency

---



- Focus:
  - Operative decision-making
  - Experienced human experts and where the stakes are high
- Human agency
  - „... AI systems should support individuals in making better, more informed choices in accordance with their goals. ...“  
(European Commission: Directorate-General for Communications Networks, Content and Technology, *Ethics guidelines for trustworthy AI*, Publications Office, 2019)

# Problems from Perspective of Work Psychology

---



- Automation capabilities exceed human capabilities (Bainbridge, 1983)
  - Humans are assigned tasks that go beyond their capabilities
  - Humans lose skills (deskilling)
- Automation complacency: Over-reliance (Parasuraman & Manzey, 2010)
  - Typical human errors: Omission error / commission error
- AI exacerbates these problems (Endsley, 2023)



# Comprehensibility: Necessary but not Sufficient

---



- Humans still over-rely to AI even if the AI is comprehensible (by explainability / interpretability) (Buçinca et al., 2021; 2024)
  - Humans tend to not engage analytically with explanations
  - Cognitive forcing does not help
  - Humans tended to accept incorrect AI recommendations, even if they would have made a better decision without AI
- When humans do not engage with AI-generated functions and do not question them, performance decreases (Dell'Acqua et al., 2023)

# From Recommendation-based AI to Supportive AI

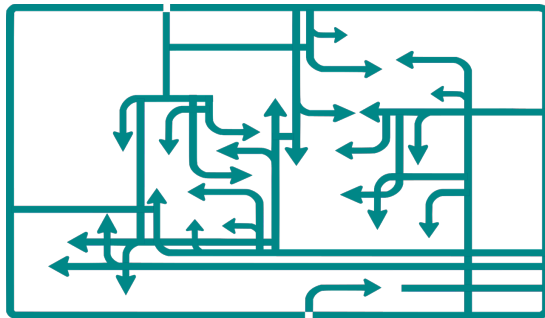


Image from Gerd Altmann auf Pixabay



Image from Mohamed Hassan auf Pixabay

## Recommendation-based AI

- Sophisticated recommendations
- Over-reliance of humans

## Supportive AI

- Supporting cognitive processes
- Augmenting human cognition

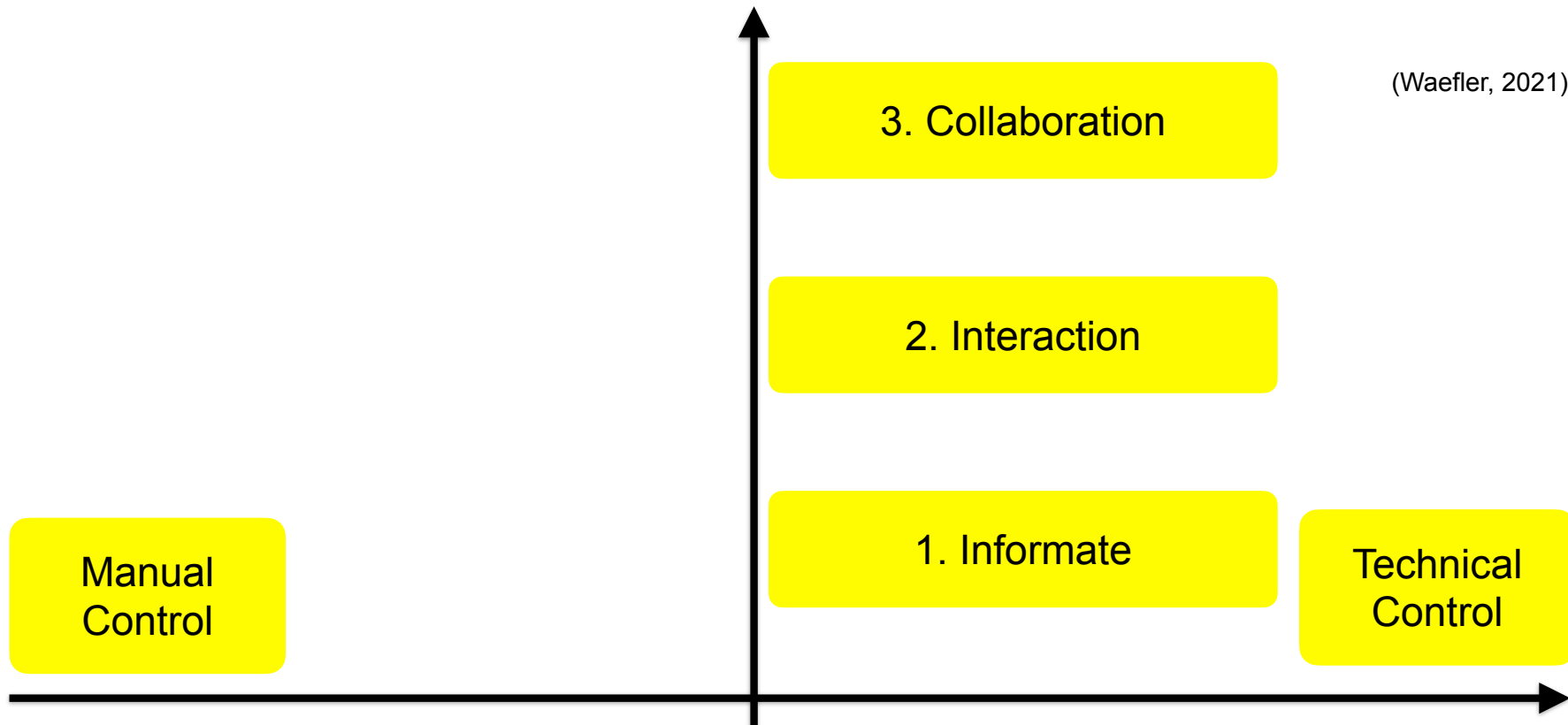


Image from OpenClipart-Vectors auf Pixabay

# Intensified Human-AI Collaboration



(Waeffler, 2021)



# Human Expertise: Tacit knowledge



Human Expertise: Pattern Recognition;  
Humans „read“ situations

Often an unconscious process

Mostly tacit: i.e. not explainable

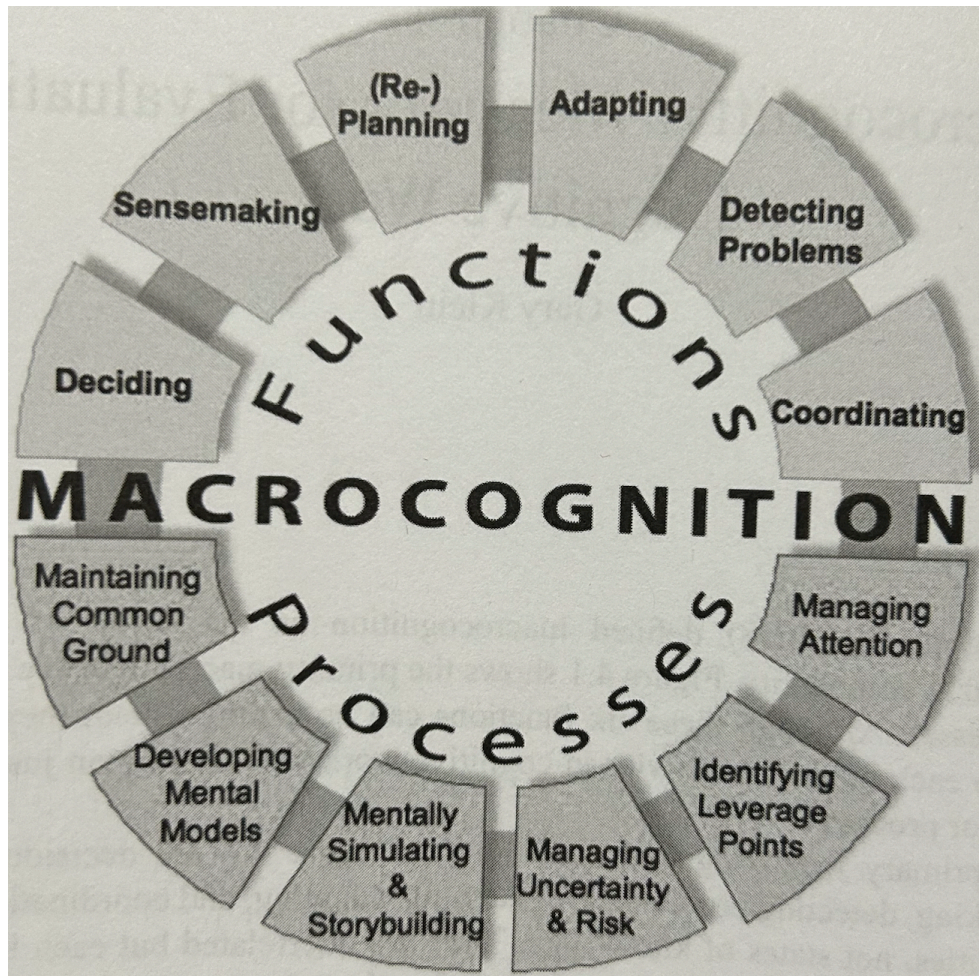
Prerequisite: Experience

- What is important about the situation?  
What needs to be taken into account?
- What needs to be done?

(Klein, 1998)

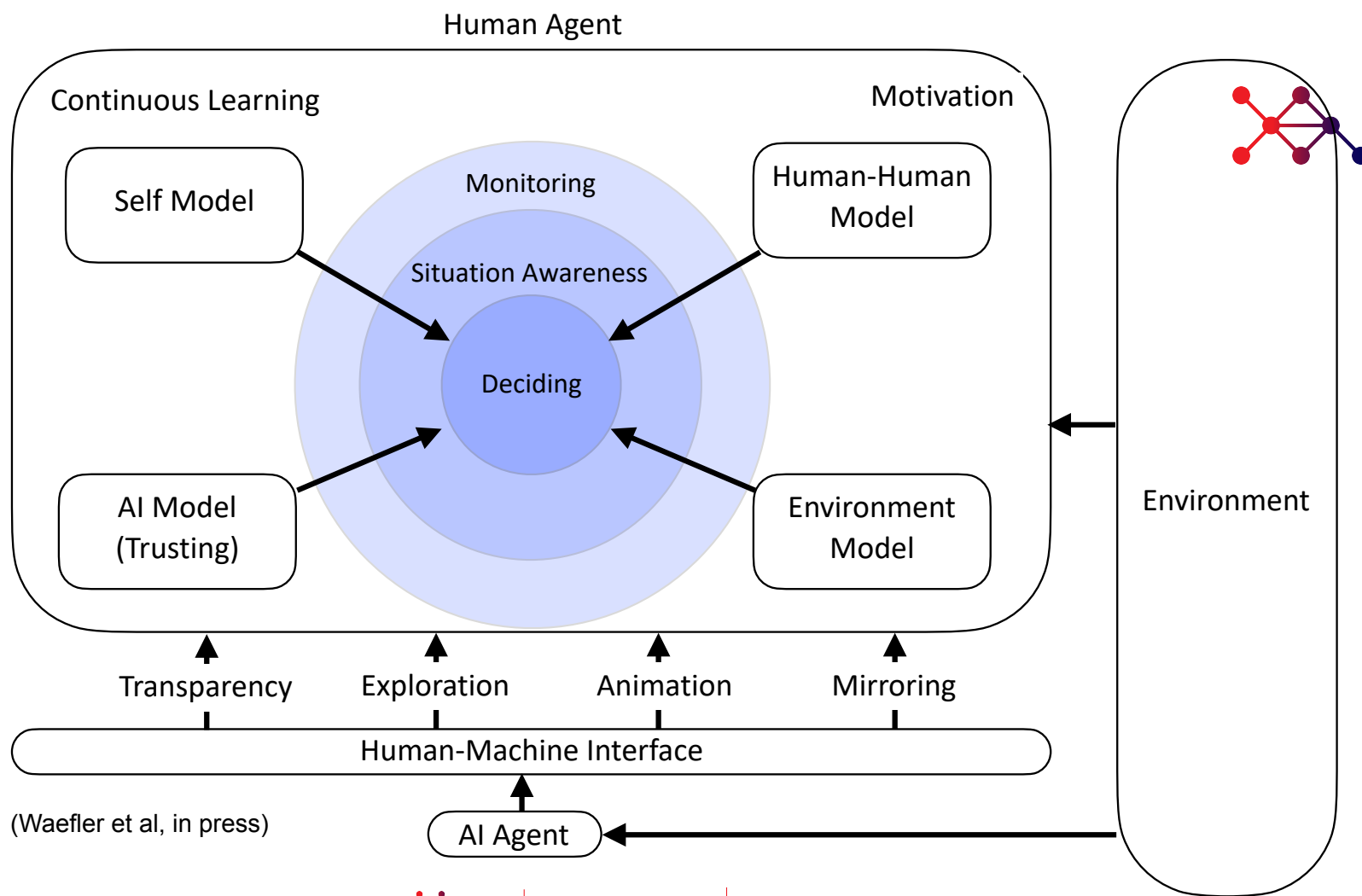


Bild von Bernd auf Pixabay



(Klein, 2018)







## Supportive-AI Explicitly Supports Human Cognitive Processes

---



- Supporting human decision-making regarding:
  - e.g. managing attention
  - e.g. comparing effects of different options for decisions
- Supporting human learning regarding:
  - e.g. identifying weak signals of emerging problems
  - e.g. building expertise regarding leverage points
- Supporting human motivation by:
  - e.g. making transparent causal relations (for experienced meaningfulness)
  - e.g. providing feedback regarding the impact of their decisions (for feedback)

# References

---



- Bainbridge, L. (1987). Ironies of Automation. In: Rasmussen, J., Duncan, K. & Leplat J. (eds.) New Technology and Human Error, pp. 271--283. Wiley, Chichester.
- Buçinca, Z., Malaya, M. B., Gajos, K. Z. (2021). To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. Proc. ACM Hum.-Comput. Interact. 5, CSCW1, Article 188 21 pages.
- 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.
- Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Kraye, L., Candelon, F. & Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013, The Wharton School Research Paper.
- Endsley, M. R. (2023). Ironies of artificial intelligence. Ergonomics, 66(11), 1656–1668.
- Klein, G. (1998). Sources of Power. How People make Decisions. The MIT Press, Cambridge MA.
- Klein, G. (2018). Macrocognitive Measures for Evaluating Cognitive Work. In E. S. Patterson & J. E. Miller (Eds.), Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams. pp. 47–64. Ashgate, Farnham.
- Parasuraman, R., Manzey, D.H. (2010) Complacency and bias in human use of automation: An attentional integration. Human Factors. 52(3), pp. 381–410.
- Waeﬂer T. (2021). Progressive Intensity of Human-Technology Teaming. Proceedings of the 5th International Virtual Conference on Human Interaction and Emerging Technologies, IHET 2021, August 27–29, 2021, France, pp. 28-36.
- Waeﬂer T. (in press). The Supportive AI Framework: From recommending to supporting. Proceedings of the 27th International Conference on Human-Computer Interaction, HCI International 2025, June 22-27, 2025, Gothenburg, Sweden.





AI4REALNET has received funding from [European Union's Horizon Europe Research and Innovation programme](#) under the Grant Agreement No 101119527

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.



[ai4realnet.eu](https://ai4realnet.eu)

