Human factors in the age of AI

Artificial Intelligence (AI) based technology is rapidly evolving and will have major implications for how humans and technology collaborate in workflows and operations. Dorthea Mathilde Kristin Vatn and Thor Myklebust use a traditional sociotechnical Human, Technology, Organization (HTO) framework to provide a theoretical introduction to the important implications for specialists working with Human Factors (HF) and suggest how the framework can be extended in the context of AI, highlighting three main implications for HF specialists in the age of AI.

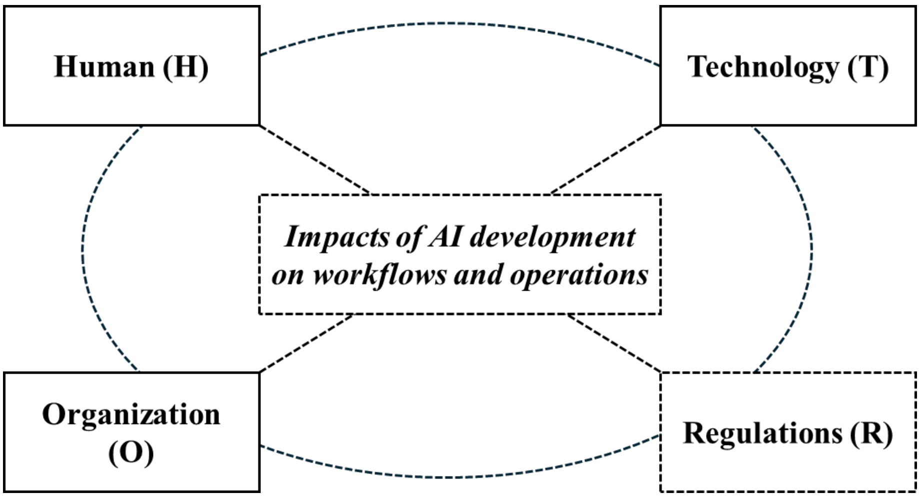
From a sociotechnical perspective, understanding how AI impacts workflows and operations requires a consideration of human, technological, as well as organizational aspects [1]. By providing equal emphasis on all aspects and the interactions amongst them, we are in a better position to understand how work is performed and the outcomes of it [2].

This idea underpins the "HTO concept" [1] highlighting that work activities can be described, analysed and understood by exploring the interactions between three sub-systems including Human (H), Technology (T), and Organization (O).

- The Human component encompasses a description of humans at four levels covering the human as a biological energy processing system, as a cognitive information processing system, as an individual with a unique history, and as a member of social groups and cultures.
- The Technology component refers to the technology itself, including specific procedures and methods tied to the use of the technology in the organizational context.
- The Organizational component refers to both the formal arrangements and informal social structures of an organization, including the structural dimensions tied to different organizational roles and hierarchies [1].

While the sociotechnical HTO concept encompass several facets that all are important to understand how AI impacts workflows and operations within organizational settings, we also believe that understanding AI in this context requires a careful consideration of regulations (R) that both affect how AI systems are developed, implemented and used in the organizational setting [3].

The figure below presents the components considered from a sociotechnical perspective when examining the impact of AI on workflows and operations, using the HTO concept [1] and incorporating regulations (R) as an additional element. We next explore how this extended HTO framework offers valuable insights into how AI is transforming organizational workflows and operational practices.



Technological aspects

Although AI systems should not be considered "novel" technologies in themselves, modern data-driven, machine learning-based systems introduce several new challenges. One first major challenge is the complexity of these systems, which often makes their inner workings opaque. As a result, many AI solutions are difficult to interpret, leaving human stakeholders unable to fully understand how these systems operate or how they arrive at their decisions [4].

This concern has led to an increasing interest in Explainable AI (XAI) as a research field [4], that from the technical perspective encompass technical tools and methods that provides insight into the inner workings of AI systems [5]. A second challenge relates to the inherently dynamic nature of modern data-driven AI systems, that are continuously evolving as they are retrained on new data. This aspect challenges the traditional view of technology development as a discrete activity, separate from its practical application in various contexts [6]. This technological challenge calls for the usefulness of complementing a model-centric view on AI development with a data-centric view on AI development.

Human aspects

While AI systems are often introduced as tools to automate tasks previously performed by humans, they frequently do not eliminate human operators but instead transform existing workflows. This shift underscores the need to consider how AI systems align with human cognitive processes, addressing concepts such as situational awareness, attention, expertise, and learning [7,8].

Accidents involving modern cars operating in autopilot mode underscore the importance of ensuring that AI-enabled features are designed to align with human cognitive capabilities. Similarly, research in the medical domain shows that reliance on AI can lead to a decline in human performance, emphasizing the risk of deskilling [9]. In addition, studies on Explainable AI (XAI) reveal that understanding how AI systems function requires careful consideration of the intended audience of explanations [4]. Although there may be a technical explanation for how an AI system operate and generate decisions, these explanations might not be understandable for certain user groups.

Organizational aspects

Beyond the technical and human aspects that need to be considered, several organizational aspects should be considered as well. While, from an organizational perspective, there are several aspects that could be discussed, we highlight two central aspects that follows from the technical and human considerations above.

"Although there may be a technical explanation for how an AI system operate and generate decisions, these explanations might not be understandable for certain user groups"

The first relate to the dynamic nature of AI systems that requires organizations to approach AI development and AI use as inseparable activities. This emphasizes the need for organizations to explore how development teams and users should work together, and how such work should be organized and structured.

The second relate to the importance of having an organizational capability to provide explanations of AI systems' operations and decisions to different stakeholders in the organizational setting [10]. Developing such a capability requires investments not only in specific XAI tools, but also people and processes [11].

Regulatory aspects

The EU Artificial Intelligence Act (AI Act, EU 2024/1689) establishes the first comprehensive and binding regulatory framework for AI, based on a risk classification that distinguishes between unacceptable, high, limited, and minimal risk systems. High-risk AI systems are subject to extensive requirements, many of which explicitly address Human Factors related issues (e.g. Article 14 – Human Oversight). The requirements cover risk management, data governance, technical documentation, human oversight, post-market monitoring, and conformity assessment prior to market access. The AI Act is embedded in the EU New Legislative Framework, linking compliance to CE marking, notified bodies, and harmonized standards, thereby integrating AI regulation into existing product safety regimes [3].

Beyond Europe, AI regulation is developing globally. Several jurisdictions, including the United States, China, Canada, the United Kingdom, Japan, and Australia, have introduced AI-related legislation, executive orders, or binding sectoral rules. These initiatives reflect increasing regulatory focus on safety, transparency, accountability, and risk management, although implemented through different legal approaches and levels of prescriptiveness.

A key implementation mechanism of the AI Act is the development of harmonized European standards. Once cited in the Official Journal of the EU, these standards provide presumption of conformity. An example is prEN 18286, a forthcoming harmonized standard defining a quality management system tailored to AI Act requirements that operationalizes Article 17 – Quality Management System by specifying organizational processes for documentation, traceability, risk management, and lifecycle governance of AI systems.

Implications of the HTOR-framework for HF Specialists

By providing examples of how AI impacts workflows and operations through the lens of the HTOR-framework we have elucidated that modern data-driven AI systems have implications covering technological, human, organizational, as well as regulatory aspects. Next, we discuss practical implications that follows our theoretical discussion above.

We need to rethink how user-centred design is performed

The dynamic nature of modern data-driven AI systems underscores the need to reconsider established approaches to user-centred design. Whereas the development of rule-based systems could adhere to traditional cycles of iterative user testing, culminating in a clear endpoint where the system is ready for deployment, determining such an endpoint is far more challenging in the context of machine learning-based systems.

From an organizational perspective, this necessitates the sustained involvement of AI expertise that collaborates closely with end users and domain specialists throughout the entire lifecycle of AI systems. These requirements, in turn, have significant implications for how work is structured and how resources are allocated to support the ongoing collaboration essential for effective AI development and integration within organizational settings.

Insight into human cognition is more relevant than ever

Both in the context of AI systems that are supposed to partly assist or take over certain work tasks, paying careful attention to human cognition when end users are introduced to new tools is increasingly important. In safety-critical systems, where human operators are retained to maintain controllability and intervene if the technical system fails or an unforeseen event occurs, it is essential to design equipment and workflows that preserve situational awareness, thereby enabling effective human oversight [3].

Furthermore, when designing tools that is supposed to assist highly skilled expertise, we need to pay careful attention to how such tools affect expertise development and learning, to avoid deskilling. This calls for the importance of using insight from fields such as psychology that can provide insight into how human learning and expertise development unfolds.

We need new standards covering human aspects

The existing set of international standards addressing Human Factors in relation to safety and AI is currently limited, representing an unresolved standardization gap.

While the next edition of IEC 61508 for functional safety introduces clearer references to AI (e.g. IEC TR 5469 and ISO/IEC TS 22440-1) and cybersecurity (e.g. the IEC 62443 series), its treatment of Human Factors remains limited and largely implicit. This is problematic as Human Factors should be addressed at a level comparable to AI and cybersecurity, with explicit normative references and lifecycle considerations in safety-related systems increasingly characterized by AI-driven behaviour.

"However, the absence of Human-Factors-specific standards relevant to the development and assessment of modern AI systems highlights a significant and unresolved standardization gap"

There are some standardization approaches that can be used, for instance ANSI/HFES 400 [12] can be used to perform a systematic assessment of human readiness and performance across the system lifecycle. However, the absence of human-factors-specific standards relevant to the development and assessment of modern AI systems highlights a significant and unresolved standardization gap.

References

[1] Karltun, A., Karltun, J., Berglund, M., & Eklund, J. (2017). HTO–A complementary ergonomics approach. Applied ergonomics, 59, 182-190.

[2] Hollnagel, E. (2017). The ETTO principle: efficiency-thoroughness trade-off: why things that go right sometimes go wrong. CRC press.

[3] Myklebust, T., Stålhane, T., & Vatn, D. M. K. (2025). The AI Act and The Agile Safety Plan (p. 156). Springer Nature.

[4] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion, 58, 82-115.

[5] Gunning, D., Vorm, E., Wang, J. Y., & Turek, M. (2021). DARPA's explainable AI (XAI) program: A retrospective. Applied AI Letters, 2(4).

[6] Waardenburg, L., & Huysman, M. (2022). From coexistence to co-creation: Blurring boundaries in the age of AI. Information and Organization, 32(4).

[7] Lee, J. D., Wickens, C. D., Liu, Y., & Boyle, L. N. (2017). Designing for people: An introduction to human factors engineering. (3ed). Create Space.

[8] Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. Human factors, 37, 85-104.

[9] Budzyń, K., Romańczyk, M., Kitala, D., Kołodziej, P., Bugajski, M., Adami, H. O., ... & Mori, Y. (2025). Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. The Lancet Gastroenterology & Hepatology, 10(10), 896-903

[10] Vatn, D. M. K. & Mikalef, P. (2025). XAI Capabilities as a lens to theorize XAI in organizations. R&D Management conference, 2025, Pisa, Italy.

[11] McKinsey & Company. (2024). Building AI trust: The key role of explainability. Accessed 6th of Jan at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/building-ai-trust-the-key-role-of-explainability

[12] ANSI/HFES 400-2021 (2021). Human Readiness Level Scale in the System Development Process. Accessed 6th of Jan at: https://www.hfes.org/Portals/0/Documents/DRAFT%20HFES%20ANSI%20HRL%20Standard%201_2_2021.pdf

Dorthea Mathilde Kristin Vatn, Research Scientist, SINTEF Digital

Dorthea is a research scientist within Human Factors, Psychology, & Information Systems at SINTEF Digital. She holds a MSc in Work and Organizational Psychology and is currently combining her role as a research scientist in SINTEF with pursuing a PhD in Information Systems. She works with questions at the intersection of people and technology, exploring how novel technologies impacts people and organizations both in a safety perspective and a business perspective.

Thor Myklebust, Senior Research Scientist, SINTEF Digital

Thor is a senior researcher in Safety and Reliability at SINTEF Digital. Since 1987 he has been involved in research, assessment and certification of products and systems. He has worked for the National Metrology Service, Aker Maritime, Nemko, and SINTEF. Thor has extensive experience in participating in several international committees and he has written four Springer books and more than 300 papers and reports.