Human Interaction and Emerging Technologies (IHIET-AI), Vol. 201, 2026, 425–434
https://doi.org/10.54941/ahfe1007196

Integrating SKM and STPA for Human-Centred Optimization of Sustainable AI-Driven Enterprise Systems

Letouzé Théodore[1], André Jean-Marc[1], Diaz-Pineda Jaime[2], and Vennin Coralie[3]

[1]IMS Laboratory UMR CNRS 5218, Bordeaux INP / ENSC, Université de Bordeaux, Talence, France
[2]Thales AVS, Mérignac, France
[3]XpLearn, Mérignac, France

ABSTRACT

The integration of artificial intelligence (AI) is reshaping corporate practices-from decision-making and innovation to production. These technologies offer performance gains but raise concerns about governance, decision reliability, security, and human AI coordination. Despite their central role in oversight, human actors remain underformalized in system design. This paper presents a conceptual framework that integrates Systemic Knowledge Management (SKM), System-Theoretic Process Analysis (STPA), and a human-centered approach to better support AI-assisted systems. It aims to move beyond a tech-centric vision by jointly modelling human, organizational, and technical dimensions. SKM structures and capitalizes knowledge, ensuring human–AI process alignment, while STPA analyses unsafe decisions, coordination failures, and control losses. The integration supports decision traceability, performance analysis, and resilience by closing the gap between knowledge modelling and system control.

Keywords: System-theoretic process analysis, Systemic knowledge management, IA, Sustainable optimization, Human factor

INTRODUCTION

The transition from Industry 4.0 to Industry 5.0 marks a paradigm shift that focused on human–AI collaboration (Verna et al., 2024). While Industry 4.0 emphasized automation, Industry 5.0 prioritizes human involvement and skill complementarity (Schroder et al., 2024; Agostinho et al., 2023). This evolution relies on AI's analytical capabilities, which support innovation but introduce challenges, especially the opacity of decision-making processes—the so-called "black box." This erodes trust and complicates human oversight (Agostinho et al., 2023; Mourtzis et al., 2022). In response, explainable AI (XAI) and strong data governance have become essential (Gadekallu et al., 2025).

However, most literature treats knowledge management and risk analysis as distinct. Few studies offer integrated frameworks that align human knowledge with algorithmic governance. This contribution proposes an

Received January 22, 2026; Revised February 21, 2026; Accepted March 8, 2026; Available online April 1, 2026

integration of Systemic Knowledge Management (SKM) and System-Theoretic Process Analysis (STPA) to secure and optimize human–AI collaboration in industrial environments.

STATE OF THE ART AND THEORETICAL FOUNDATIONS

AI in Industrial and Company Systems: Opportunities and Limitations (Focus on Transparency)

Artificial intelligence (AI) is profoundly transforming industrial and organizational systems by redefining strategic decision-making processes, innovation, design, supply chain management, and production. The opportunities associated with its implementation include process optimization, cost reduction, quality improvement, product customization, and accelerated innovation (Haefner et al., 2021). By enhancing the speed, accuracy, and objectivity of managerial decisions through the massive use of real-time data, the integration of AI is revolutionizing strategic business management (Mohamed et al., 2025). Thus, without entirely replacing human judgment, it acts as a tool for augmenting decision-making capabilities, reducing cognitive biases, and optimizing resource allocation in volatile and complex environments. However, effective and sustainable adoption depends on rigorous governance that can reconcile the technical capabilities of AI with issues of trust, data quality and transparency, as well as ethical and organizational requirements (Narne, 2024). In this context, transparency is critical for trust, regulatory compliance, and governance. Explainable AI is thus essential in high-stakes sectors where algorithmic decisions must be interpretable and auditable (Gadekallu et al., 2025; Richey et al., 2023). The implementation of AI in industrial and business systems cannot therefore be limited to a technological approach. It requires appropriate governance, risk management policies, audit mechanisms, and enhanced compliance, with a particular focus on ethics, data protection, and skills development (McCormack et al., 2025). Thus, the organizational and social acceptability of AI largely relies on the ability of organizations to make these systems transparent, explainable, and trustworthy (Richey et al., 2023; Gadekallu et al., 2025). Ultimately, while While AI enhances industrial performance, its deployment depends on solving key issues: transparency, data quality, and human-technical alignment (Ahmed et al., 2022; Agostinho et al., 2023; Gadekallu et al., 2025).

Systemic Knowledge Management: Definition and Challenges Associated With Capitalizing on Tacit and Explicit Knowledge

The distinguishing between tacit and explicit knowledge is a central pillar of knowledge management theories. Tacit knowledge refers to experiential know-how that is contextualized and incorporated into individual practices, while explicit knowledge is formalized, codified, and transferable through documents, procedures, or models (Nonaka, 1994). From this perspective, Systemic Knowledge Management (SKM) is not limited to knowledge preservation, but aims to align its production, structuring, and dissemination with the operational and strategic objectives of the organization. The operationalization of SKM relies on the implementation of structured mechanisms for capturing, formalizing, and validating knowledge derived from actual work. Previous work highlights the structuring role of feedback processes in transforming tacit, situated, and experiential knowledge into reusable explicit knowledge (Kamsu Foguem et al., 2008). These mechanisms are based on formalized steps: description of the event, contextualization, analysis, and formulation of solutions, promoting the externalization of individual knowledge. In order to ensure that this knowledge is shared and reused, it must be structured using semantic tools, such as domain ontologies. Governance mechanisms, including multidisciplinary committees, ensure the validation and legitimization of the knowledge produced, guaranteeing its consistency with standards and organizational strategy (Nonaka, 1994). Beyond a heritage-based approach, SKM is a lever for operational performance in complex industrial environments (Fugate, Stank, and Mentzer, 2009). By highlighting the rules, heuristics, and compromises that structure actual work practices, it helps reduce variability in operating procedures and dependence on individual expertise (Stürzebecher et al., 2025). Several studies show that structured knowledge management approaches can reduce resolution times, non-quality costs, and operational disruptions, particularly in contexts marked by turnover, technical complexity, or system criticality (Kamsu Foguem et al. 2008; Tan et al., 2015). SKM thus acts as an organizational infrastructure that stabilizes action and decision-making at the operational level (Fugate et al., 2009).

The contribution of Systemic Knowledge Management (SKM) becomes particularly critical in industrial systems that are increasingly supported by artificial intelligence devices (Carayannis et al., 2021; Samtani et al., 2023). When decisions are co-produced by human operators and algorithmic systems, operational performance depends heavily on the alignment between the mental models of the actors, organizational rules, and AI-generated recommendations. In this context, explicit practices, reasoning, and arbitration criteria provided by SKM are a prerequisite for analyzing and governing human-AI interactions within complex socio-technical systems (Carayannis et al., 2021; Zhang et al., 2025). This necessity justifies the use of systemic approaches such as System-Theoretic Process Analysis (STPA), which view safety and performance as emerging control problems within socio-technical systems. They enable the holistic identification of loss scenarios and control constraints related to interactions between human, organizational, and technical components (Karevan & Nadeau, 2024; Patriarca, Chatzimichailidou, Karanikas, & Di Gravio, 2022; Sadeghi & Goerlandt, 2023).

Limitations of Traditional Risk Analysis Approaches and the Contribution of STPA–SKM Integration

In their traditional form, industrial risk analysis methods such as FMEA or HAZOP (Levinson, 2012) remain focused on component failures and essentially linear causal chains. This focus makes them ill-suited to complex, highly coupled and largely software-based socio-technical systems, which are characteristic of AI applications (Abdulkhaleq, Wagner & Leveson, 2015; Patriarca, Chatzimichailidou, Karanikas & Di Gravio, 2022). Conversely, the STPA approach, based on the STAMP (Systems-Theoretic Accident Model and Processes) model, views safety as a control issue (Chen et al., 2025). Accidents are then interpreted as the consequence of a control structure that is unable to properly enforce safety constraints due to inadequate interactions, human decisions, software or organisational errors, even in the absence of hardware failure (Leveson, 2012, 2015; Rising & Leveson, 2018). STPA therefore focuses on analysing feedback loops, control actions and controller process models in order to identify Unsafe Control Actions and define safety constraints that can be traced back to design and operational requirements (Leveson, 2012; Nakhal, Patriarca, De Carlo & Leoni, 2023; Pawlicki et al., 2016).

In AI systems, where behaviours emerge from interactions between algorithms, human operators and organisation, this systemic approach makes it possible to move beyond a vision centred on ‘component failure’ and to consider security as the maintenance of control constraints in dynamic environments. However, in the absence of mechanisms dedicated to capitalising on results, the lessons learned from STPA often remain isolated, difficult to reuse and poorly integrated into operational processes. This limitation highlights the need to link STPA with knowledge management approaches capable of structuring, tracking and developing the knowledge gained from safety analysis, thus paving the way for an integrated framework combining SKM and STPA. The integration of STPA with knowledge management (KM) approaches is therefore essential for organising, documenting and updating the knowledge generated by these analyses, and for embedding it sustainably in industrial practices. With this in mind, the MASK method (Matta et al., 2002) offers real-world models that can improve STPA process models and enhance the relevance of safety analyses (Benmahamed et al., 2005).

Proposals for an Integrative Vision of Industrial Performance

Industrial performance within socio-technical systems incorporating artificial intelligence relies on the ability to simultaneously control the knowledge mobilised by actors and the control mechanisms governing decisions and interactions. While AI systems help improve the efficiency, quality and responsiveness of industrial processes, their integration profoundly transforms decision-making modes and the balance between human actors and automated processes (Ahmed et al., 2022; Bahoo et al., 2023). In this context, performance cannot be dissociated from the conditions of understanding, appropriation and governance of algorithmic recommendations. Our proposal is based on the hypothesis that the combination of systemic knowledge management (SKM) and system-theoretic process analysis (STPA) can jointly improve the safety, decision-making robustness and operational performance of industrial systems.

The link between SKM and STPA, particularly through the MASK method, is based on a structural correspondence between knowledge models and the constituent elements of control analysis. The activity models and reference system developed using MASK are used to define the scope of the system and identify the actors, flows and resources involved, thus providing the basis for defining the control structure in STPA (Aries et al., 2008; Pereira et al., 2019). Task and concept models formalise the reasoning, problem-solving strategies and state variables used by human or automated controllers, enriching the analysis of process models at the heart of STPA. In addition, phenomenon, history and lineage models provide a framework for examining controlled processes, environmental disturbances and system evolutions, facilitating the identification of causal scenarios and the traceability of safety constraints. This complementarity allows control analysis to be anchored in real practices and organisational knowledge, rather than remaining at a purely abstract level.

In this context, Figure 1 illustrates an integrative framework combining systemic knowledge management (SKM/MASK) and systemic process analysis (STPA) to support the governance, safety and operational performance of socio-technical systems incorporating artificial intelligence. The actual work within the real socio-technical system (human actors, AI-based recommendations and operational constraints) is the main source of knowledge and feedback. STPA guides SKM activity. SKM/MASK captures and formalises this experience (REX) in activity, task and concept models, which are transformed into formalised knowledge (scenarios and constraints) feeding into the STPA assessment process. STPA analyses control structures, process models and unsafe control actions to derive safety and performance constraints. These constraints inform the organisational governance of human-AI decision-making and operational performance, while continuous improvement loops ensure iterative learning and adaptation through incident analysis and organisational feedback.

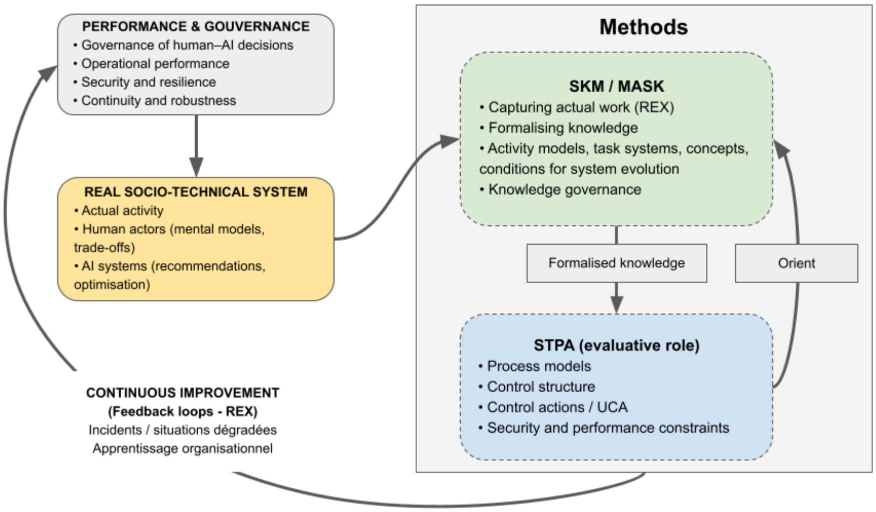

Figure 1: Integrated framework for organisational performance and governance in AI-based socio-technical systems.

To illustrate the contribution of the proposed framework, let us consider a typical scenario of human-AI interaction. In a context of production assisted by a recommendation system, an operator applies an automatically generated instruction without questioning its validity, due to overconfidence in the system and the absence of feedback on the actual status of the process. The SKM provides tools for formalisation (modelling of business rules, identification of weak signals and the arbitration criteria used by experienced operators); these elements are then integrated as explicit constraints in the STPA analysis. The STPA analysis identifies this situation as an Unsafe Control Action, resulting from an incomplete process model. The causal scenario then highlights a lack of explainability, insufficient training and a lack of escalation rules. Corrective actions can then be implemented, including technical adjustments (interfaces, interactions, space ergonomics, etc.), the implementation of procedures (standardisation, human validation) and targeted training. This combination helps to reduce recurring errors, speed up decision-making in disruptive situations and maintain operational continuity. The results of this theoretical analysis show that integrating SKM and STPA is an effective lever for strengthening the operational performance and resilience of AI-based industrial systems. However, implementing such an integrative framework requires a significant initial investment in time and skills, as well as a high level of organisational maturity in terms of knowledge management and safety culture. Cooperation between business experts, systems engineers and AI specialists remains a critical factor, as does the ability to maintain and update knowledge produced in rapidly changing environments. These constraints can hinder the adoption of the framework, particularly in organisations with limited resources or highly compartmentalised structures. Thus, the combination of SKM and STPA appears to be a structuring approach for consolidating the reliability, safety and performance of industrial systems incorporating artificial intelligence. While AI optimises planning, maintenance and quality (Ahmed et al., 2022; Bahoo et al., 2023), SKM helps to preserve knowledge, standardise best practices and accelerate problem solving (Nonaka, 1994; Kamsu Foguem et al., 2008). By combining these contributions with STPA's systemic analysis, the proposed framework makes it possible to anticipate the pitfalls of excessive reliance on automated systems and to strengthen the robustness of operational decisions.

CONCLUSION

The relationship between KM and STPA tends to improve the operational performance and resilience of the socio-technical system by taking into account the processing of interactions and knowledge. AI improves efficiency, quality, planning and maintenance (Bahoo et al., 2023; Ahmed et al., 2022). SKM reduces knowledge loss, accelerates problem solving and standardises best practices (Nonaka, 1994; Kamsu Foguem et al., 2008). Our theoretical analysis suggests that combining Systemic Knowledge Management and STPA is a lever for improving the operational performance and robustness of industrial systems incorporating AI. Indeed, AI supports process and management optimisation, but its opacity undermines trust and ownership, particularly when decisions are co-produced by human actors and algorithmic systems (Agostinho et al., 2023; Gadekallu et al., 2025). Meanwhile, SKM strengthens operational continuity by making explicit the rules, trade-offs and feedback from actual work, facilitating their dissemination and reuse (Nonaka, 1994; Kamsu Foguem et al., 2008). Combined with STPA, this capitalisation feeds into the analysis of process models and enables more systematic identification of loss scenarios linked to interactions and feedback deficits (Leveson, 2012; Patriarca et al., 2022). The implementation of an integrative SKM–STPA framework remains challenging, however, as it requires significant resources, multidisciplinary coordination, and sufficient organisational maturity to maintain the traceability of knowledge and control constraints over time (Bougoulia & Glykas, 2022; Raeisdanaei, Kim, Liao, & Kochhar, 2025). Without explicit governance and without management and integration of knowledge, there is a risk of producing ad hoc analyses that are poorly integrated into operations and insufficiently updated when the system evolves (Hitziger et al., 2019; Hoffmann, Pohl, & Hering, 2017; Oliver et al., 2021; Scholz et al., 2024). Our proposal combines knowledge capitalisation and systemic control analysis to strengthen the governance of human-AI interactions in an Industry 5.0 context. The perspectives concern the operationalisation of the framework (tools, REX routines, traceability), as well as alignment with the growing requirements for transparency, auditability and certification of industrial AI systems.

ACKNOWLEDGMENT

This study builds on previous work carried out as part of the GIS Albatros 'STAP' project funded by Thales.

REFERENCES

Abdulkhaleq, A., Wagner, S., & Leveson, N. (2015). A comprehensive safety engineering approach for software-intensive systems based on STPA. Procedia Engineering, 128, 2-11. https://doi.org/10.1016/j.proeng.2015.11.498

Agostinho, C., Dikopoulou, Z., Lavasa, E., Perakis, K., Pitsios, S., Branco, R., Reji, S., Hetterich, J., Biliri, E., Lampathaki, F., Rodríguez Del Rey, S., & Gkolemis, V. (2023). Explainability as the key ingredient for AI adoption in Industry 5.0 settings. Frontiers in Artificial Intelligence, 6, 1264372. https://doi.org/10.3389/frai.2023.1264372

Ahmed, I., Jeon, G., & Piccialli, F. (2022). From Artificial Intelligence to Explainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where. IEEE Transactions on Industrial Informatics, 18(8), 5031–5042. https://doi.org/10.1109/TII.2022.3146552

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable Artificial Intelligence (XAI) : What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion, 99, 101805. https://doi.org/10.1016/j.inffus.2023.101805

Aries, S., Le Blanc, B., & Ermine, J. L. (2008). MASK: une méthode d'ingénierie des connaissances pour l'analyse et la structuration des connaissances. Management et ingénierie des connaissances: modèles et méthodes.

Bahoo, S., Cucculelli, M., & Qamar, D. (2023). Artificial intelligence and corporate innovation: A review and research agenda. Technological Forecasting and Social Change, 188, 122264. https://doi.org/10.1016/j.techfore.2022.122264

Benmahamed, D., Ermine, J.-L., & Tchounikine, P. (2005). From MASK Knowledge Management Methodology to Learning Activities Described with IMS – LD. In K.-D. Althoff, A. Dengel, R. Bergmann, M. Nick, & T. Roth-Berghofer (Éds.), Professional Knowledge Management (Vol. 3782, pp. 165–175). Springer Berlin Heidelberg. https://doi.org/10.1007/11590019_20

Bougoulia, E., & Glykas, M. (2023). Knowledge management maturity assessment frameworks: A proposed holistic approach. Knowledge and Process Management, 30(4), 355-386. https://doi.org/10.1002/kpm.1731

Carayannis, E. G., Grigoroudis, E., Del Giudice, M., Della Peruta, M. R., & Sindakis, S. (2021). An exploration of contemporary organizational artifacts and routines in a sustainable excellence context. Journal of Knowledge Management, 25(6), 1446–1466. https://doi.org/10.1108/JKM-02-2020-0108

Chen, S., El Badaoui, H., James Elizebeth, M., Nakashima, T., Khastgir, S., & Jennings, P. (2025). A scalable framework for the management of STPA requirements: A case study on eVTOL operations. ArXiv.

Fugate, B., Stank, T., & Mentzer, J. (2009). Linking improved knowledge management to operational and organizational performance. Journal of Operations Management, 27, 247–264. https://doi.org/10.1016/j.jom.2008.09.003

Gadekallu, T. R., Kumar Reddy Maddikunta, P., Boopathy, P., Deepa, N., Chengoden, R., Victor, N., Wang, W., Wang, W., Zhu, Y., & Dev, K. (2025). XAI for Industry 5.0—Concepts, Opportunities, Challenges, and Future Directions. IEEE Open Journal of the Communications Society, 6, 2706–2729. https://doi.org/10.1109/OJCOMS.2024.3473891

Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. Technological Forecasting and Social Change, 162, 120392. https://doi.org/10.1016/j.techfore.2020.120392

Hitziger, M., Aragrande, M., Berezowski, J. A., Canali, M., Del Rio Vilas, V., Hoffmann, S., ... & Rüegg, S. R. (2019). EVOlvINC: evaluating knowledge integration capacity in multistakeholder governance. Ecology and Society, 24(2), art36. https://doi.org/10.5751/es-10935-240236

Hoffmann, S., Pohl, C., & Hering, J. G. (2017). Exploring transdisciplinary integration within a large research program: Empirical lessons from four thematic synthesis processes. Research Policy, 46(3), 678–692.https://doi.org/10.1016/j.respol.2017.01.004

Jackson, P., Mesquita, A., & Silva, M. M. (2024). Human-centered artificial intelligence in industrial decision-making systems. Computers & Industrial Engineering, 188, 109847. https://doi.org/10.1016/j.cie.2024.109847

Kamsu Foguem, B., Coudert, T., Béler, C., & Geneste, L. (2008). Knowledge formalization in experience feedback processes: An ontology-based approach. Computers in Industry, 59(7), 694-710. https://doi.org/10.1016/j.compind.2007.12.014

Leveson, N. G. (2016). Engineering a safer world: Systems thinking applied to safety (p. 560). The MIT Press.

Leveson, N. (2015). A systems approach to risk management through leading safety indicators. Reliability engineering & system safety, 136, 17–34. https://doi.org/10.1016/j.ress.2014.10.008.

Matta, N., Ermine, J. L., Aubertin, G., & Trivin, J. Y. (2002). Knowledge Capitalization with a knowledge engineering approach: the MASK method. In Knowledge management and organizational memories (pp. 17–28). Boston, MA: Springer US.

McCormack, L., Bendechache, M., Lewis, D., & Huyskes, D. (2025). Trust and transparency in AI: industry voices on data, ethics, and compliance. AI & SOCIETY, 1–29.

Mohammed, I. A., Pandey, P., & Sruthi, S. (2025). The Impact Of AI On Strategic Decision Making In Modern Management. European Economic Letters (EEL), 15(3), 3770–3782.

Mourtzis, D., Angelopoulos, J., & Panopoulos, N. (2022). A Literature Review of the Challenges and Opportunities of the Transition from Industry 4.0 to Society 5.0. Energies, 15(17), 6276. https://doi.org/10.3390/en15176276

Nakhal, A. J., Patriarca, R., De Carlo, F., & Leoni, L. (2023). A system-theoretic fuzzy analysis (STheFA) for systemic safety assessment. Process Safety and Environmental Protection, 174, 575–589. https://doi.org/10.1016/j.psep.2023.07.014

Narne, S., Adedoja, T., Mohan, M., & Ayyalasomayajula, T. (2024). AI-driven decision support systems in management: Enhancing strategic planning and execution. International journal on recent and innovation trends in computing and communication, 12(1), 268–276.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. Organization Science, 5(1), 14–37.

Oliver, T., Benini, L., Borja, Á., Dupont, C., Doherty, B., ... Tarrason, L. (2021). Knowledge architecture for the wise governance of sustainability transitions. Environmental Science & Policy.

Patriarca, R., Chatzimichailidou, M. M., Karanikas, N., & Di Gravio, G. (2022). The past and present of System-Theoretic Accident Model and Processes (STAMP) and its associated techniques: A scoping review. Safety Science, 146, 105566. https://doi.org/10.1016/j.ssci.2021.105566

Pawlicki, T., Samost, A., Brown, D., Manger, R. P., Kim, G.-Y., & Leveson, N. (2016). Application of systems and control theory-based hazard analysis to radiation oncology. Medical Physics, 43(3), 1514–1530.

Raeisdanaei, A., Kim, J., Liao, M., & Kochhar, S. (2025). An LLM-integrated framework for completion, management, and tracing of STPA. ArXiv.

Richey, R. G., Chowdhury, S., Davis-Sramek, B., Giannakis, M., & Dwivedi, Y. K. (2023). Artificial intelligence in logistics and supply chain management: A primer and roadmap for research. Journal of Business Logistics, 44(4), 532–549. https://doi.org/10.1111/jbl.12364

Rising, J. M., & Leveson, N. (2018). Systems-Theoretic Process Analysis of space launch vehicles. Journal of Space Safety Engineering, 5(3), 161–170.

Samtani, S., Chinnam, R. B., & Murthy, C. A. (2023). Human–AI collaboration in industrial decision-making: A knowledge-based systems perspective. Knowledge-Based Systems, 266, 110357. https://doi.org/10.1016/j.knosys.2023.110357

Scholz, R. W., Zscheischler, J., Köckler, H., Czichos, R., Hofmann, K.-M., & Sindermann, C. (2024). Transdisciplinary knowledge integration – Part I: Theoretical foundations and an organizational structure. Technological Forecasting and Social Change.

Schröder, A. J., Cuypers, M., & Götting, A. (2024). From Industry 4.0 to Industry 5.0: The triple transition digital, green and social. In Industry 4.0 and the Road to Sustainable Steelmaking in Europe: Recasting the Future (pp. 35–51). Cham: Springer International Publishing.

Stürzebecher, P., Bayrhammer, E., & Brauckmann, N. (2025). From Tacit to Explicit: Using Live Documentation and Feedback Loops to Facilitate Knowledge Transfer. European Conference on Knowledge Management. https://doi.org/10.34190/eckm.26.2.3839