SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2025
ORGANIZATION AND MANAGEMENT SERIES NO. 233

INTEGRATING SYSTEMATIC INVENTIVE THINKING WITH RESPONSIBLE INNOVATION USING AI AS CO-CREATIVE PARTNER

Paweł WAWRZAŁA

Silesian University of Technology, Faculty of Organization and Management; Pawel.Wawrzala@polsl.pl,
ORCID: 0000-0001-9974-7689

Purpose: This paper addresses the critical gap in contemporary innovation management where traditional systematic methodologies like Systematic Inventive Thinking (SIT) operate independently from responsible innovation requirements. The research aims to develop an integrated framework that combines SIT methodology with sustainability criteria while leveraging Large Language Models as co-creative partners throughout the innovation process.
Design/methodology/approach: The study employs a conceptual framework development approach, integrating comprehensive literature review across innovation management theories, responsible innovation frameworks, and human-AI collaboration research. The methodology involves designing the RWI Framework architecture with AI-Personas system, prompt engineering strategies, and repository structures. The approach synthesizes established theories including stage-gate processes, design thinking, and open innovation paradigms with emerging AI capabilities and sustainability frameworks.
Findings: The research presents the RWI Framework featuring six specialized AI-Personas that facilitate systematic integration of SIT techniques with responsibility evaluation. Key findings include the identification of AI-Coach as a knowledge transfer mechanism bridging competency gaps in sustainability, the development of multi-dimensional evaluation criteria based on established frameworks (ESG, GRI, CSRD), and the conceptualization of hybrid human-AI collaborative workflows that maintain creative efficiency while ensuring systemic responsibility assessment.
Research limitations/implications: The framework inherits current LLM limitations including hallucination risks and output variability. Future research should focus on empirical validation through case studies across diverse organizational contexts and controlled experiments comparing traditional SIT processes with RWI-enhanced approaches. The study suggests investigating knowledge transfer mechanisms and factors determining human-AI partnership effectiveness.
Practical implications: Implementation promises operational efficiency gains through automated documentation, enhanced innovation quality via early risk identification, and reduced regulatory and reputational risks. Organizations can expect development of new hybrid competencies and specialized roles such as AI Innovation Facilitators while building systematic approaches to responsible innovation.

Social implications: The framework supports integration of UN Sustainable Development Goals and ESG criteria into organizational innovation processes, potentially accelerating sustainable solution development and enhancing corporate social responsibility practices across industries.
Originality/value: This represents the first systematic framework integrating SIT methodology with AI-powered responsible innovation evaluation, addressing the identified competency gap in sustainability knowledge while pioneering co-creative human-AI partnerships in systematic innovation processes.
Keywords: Systematic Inventive Thinking, SIT, Responsible Innovation, Human-AI Collaboration, Sustainability Frameworks, Innovation Management.
Category of the paper: Conceptual paper, viewpoint.

1. Introduction

Contemporary organizations face multidimensional innovation challenges that extend far beyond traditional technological dilemmas. On one hand, there exists a need for systematic methods of generating breakthrough solutions capable of ensuring competitive advantage. On the other hand, organizations must consider growing requirements regarding social and environmental responsibility, which determine long-term market success and access to investment capital.

Traditional innovation approaches often lead to solutions that are business-effective in the short term but not necessarily sustainable in the systemic context. Simultaneously, emerging capabilities of generative artificial intelligence (GenAI) open new horizons in supporting creative processes but carry specific technical and methodological risks (Gerlich, 2025; Sidhpurwala, 2024).

This paper presents the RWI Framework (Responsible World Integration), a structured approach that combines Systematic Inventive Thinking (SIT) methodology with external responsibility criteria. The framework utilizes Large Language Models (LLM) capabilities not only for automating evaluative processes but primarily as co-creative partners throughout the entire innovation cycle (Wang et al., 2025; Viswanathan et al., 2025). This solution allows organizations to maintain efficiency and systematicity in the innovation process while supporting compliance with global sustainable development standards and building competencies in responsible innovation.

2. Research Problem

Contemporary innovation processes are characterized by three fundamental challenges that have not been addressed in an integrated manner to date. First, traditional innovation methodologies such as SIT focus on technical and business efficiency while omitting systemic aspects of social and environmental responsibility (Goldenberg, Mazursky, 2002). Second, innovation teams operate under pressure from business objectives, often lacking specialized knowledge in sustainable development frameworks, leading to a **competency gap** (Venn et al., 2022; Goropečnik et al., 2025). Third, growing capabilities of LLM technology require new models of human-AI collaboration that extend beyond traditional tool-based approaches (Stanford HAI, 2025; Koivisto, Grassini, 2023).

Previous research has demonstrated that simple integration of SIT project teams with LLM is possible (Wawrzała, 2024), establishing the foundation for this continued investigation. The RWI Framework addresses these challenges by proposing systematic integration of structured innovation methods with systemic responsibility knowledge transfer mechanisms, utilizing LLM as co-creative partners throughout the innovation process.

3. Literature Review and Theoretical Positioning

3.1. Innovation Methodologies within Established Theories

The RWI Framework positions itself within innovation theory as a complement and hybridization of existing approaches. Unlike **stage-gate processes**, which focus on sequential business risk control (Cooper, 2010), the RWI Framework introduces parallel systemic responsibility evaluation at every stage. While **design thinking** emphasizes human-centered innovation, the RWI Framework extends this perspective to **system-centered responsibility**, considering a broader ecosystem of stakeholders (Wawrzała, 2022).

Compared to **open innovation paradigms** (Chesbrough, 2003), the RWI Framework operates in a hybrid model, combining the “Closed World” assumption of SIT with “open evaluation” against external criteria. This represents a novel approach that maintains advantages of structured idea generation methods while integrating them with systemic responsibility perspective.

3.2. SIT Methodology Assumptions and Systemic Limitations

Systematic Inventive Thinking (SIT) represents one of the established systematic innovation methodologies, based on five fundamental techniques: subtraction, multiplication, division, task unification, and attribute dependency (Goldenberg et al., 1999; Goldenberg, Mazursky, 2002).

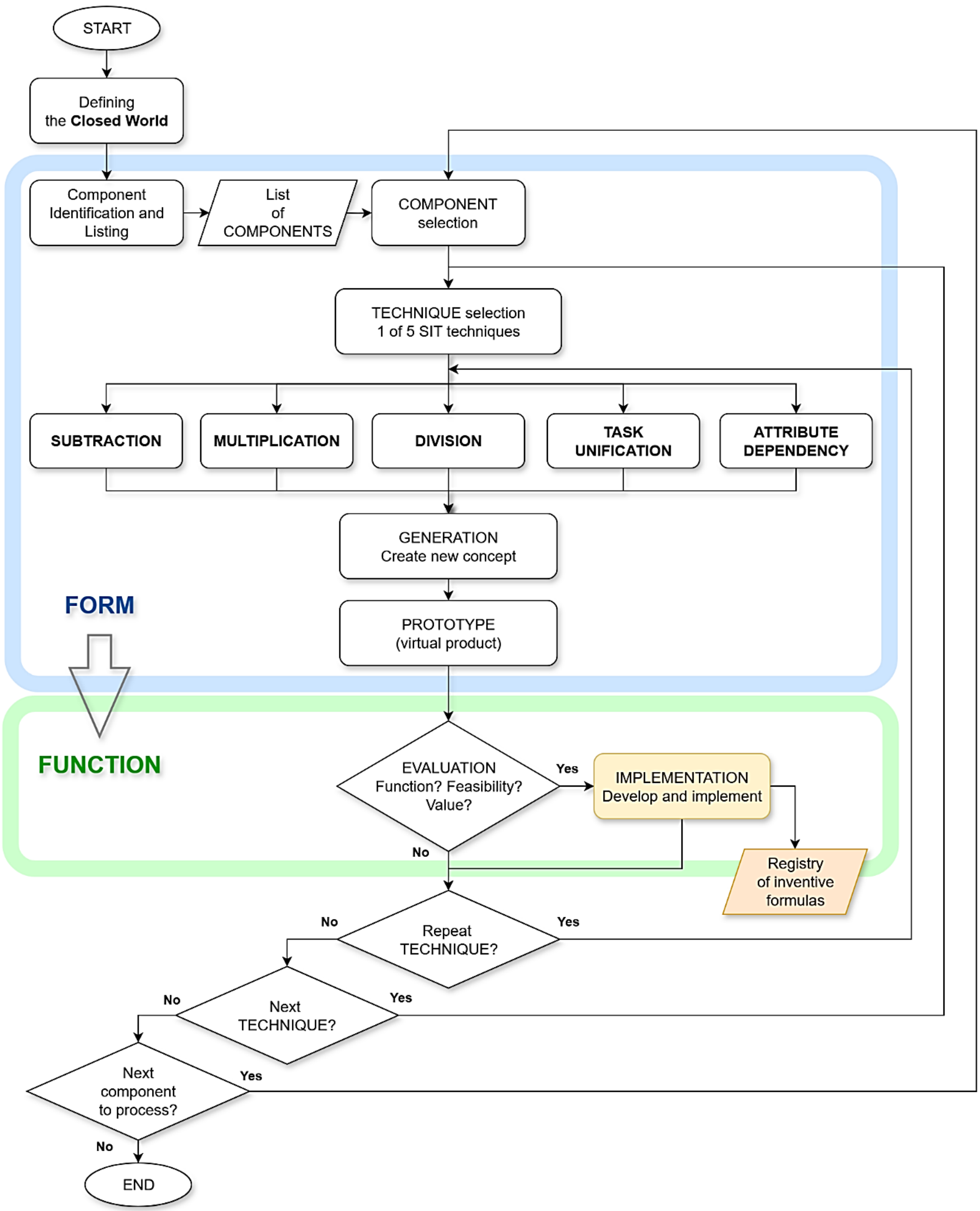

Figure 1. The complete diagram of Systematic Inventive Thinking (SIT) methodology.

Source: Own elaboration.

Figure 1 presents the complete diagram of SIT methodology, illustrating the comprehensive process from “Closed World” definition to implementation decision regarding the final result.

The methodology encompasses three distinct iteration groups that operate systematically throughout the innovation process. The component iteration loop involves systematic examination of all system elements, with decisions made regarding which components should be subjected to SIT techniques. The technique loop applies one or more of the five classical SIT techniques to the selected component. The technique iteration loop enables multiple repetitions of the chosen technique, generating successive concepts and virtual prototypes while conducting evaluations of value, functionality, and feasibility.

The SIT method demonstrates considerable capability for methodically generating substantial numbers of novel solutions. Critical factors in this process include the arbitrary division into components, the strategic decisions regarding which components to subject to technique application, and the establishment of appropriate evaluation systems. This systematic approach ensures comprehensive exploration of the innovation space while maintaining structured progression toward viable solutions.

This methodology relies on the “Closed World” principle, which limits the solution search space to elements already existing in the system or its immediate environment. This conscious imposition of constraints aims to overcome cognitive functional fixedness. The second pillar of SIT philosophy is the “Function Follows Form” principle, which reverses traditional product development logic (Finke, 1992).

Structured methods such as SIT, have proven their effectiveness in specific domains. They are particularly effective in generating product and process innovations in B2B, technological, medical, and manufacturing industries (Goldenberg et al., 2001). However, this paradigm possesses a fundamental limitation: its intensive focus on technical logic may lead to ignoring key systemic factors related to social and environmental responsibility.

While the “Closed World” assumption increases creative process efficiency, sustainable development problems inherently have an “open system” character. The RWI Framework resolves this contradiction through a two-stage approach: SIT is used for effective idea generation within the defined system, and subsequently these ideas are evaluated in the context of external criteria.

3.3. Sustainable Development Challenges and Competency Gap

Organizations currently operate in an environment of growing stakeholder expectations regarding social and environmental responsibility. The UN Sustainable Development Goals (SDGs), ESG regulations (Environmental, Social, Governance), and increased consumer awareness require systematic consideration of these aspects in innovation processes.

Innovation teams often operate under pressure from business objectives that do not consider responsibility dimensions, and their members may lack specialized knowledge in sustainable development frameworks. This creates a **competency gap**: teams are asked for innovations compliant with principles they do not fully understand (Venn et al., 2022). Each innovation process participant typically holds different views on what should be done or remains passive, often not revealing their thoughts, as these ideas may be mere seeds rather than fully formed concepts. Moreover, each participant may harbor different concerns regarding psychological safety, which further compounds communication difficulties (Wawrzała, 2022).

3.4. Specific Challenges Related to LLM Implementation

Integration of LLM technology into innovation processes carries a set of specific challenges. Language models, as probabilistic systems, operate on statistical patterns without factual verification mechanisms, resulting in generation of incomplete or false information (hallucinations) (Sidhpurwala, 2024). Users may exhibit tendencies toward anthropomorphization and excessive trust in generated content, which reduces critical vigilance (Deshpande et al., 2023). A significant challenge is the risk of cognitive offloading, where excessive reliance on AI may lead to atrophy of key cognitive skills and homogenization of ideas (Gerlich, 2025).

4. Methodology - RWI Framework Architecture

4.1. Human-AI Co-creative Partnership Concept

The RWI Framework proposes reformulating the relationship between human and AI, transitioning from a tool-based model to a co-creative partnership model. In this model, interaction always occurs along the **Team ↔ AI-Persona** line. The system operates based on architecture that preserves SIT methodology advantages while enriching the process with systemic responsibility perspective.

The process consists of three levels:

1. Defining the “Closed World” for the considered problem and conducting initial component identification, utilizing LLM as a partner possessing domain knowledge in systemic responsibility.
2. Generating solutions using SIT techniques, with enrichment of the ideation process through AI as a generative partner.
3. Systemic validation using traditional technical and business criteria as well as utilizing evaluation regarding systemic responsibility aspects, also with LLM support.

4.2. RWI Framework System Components

The system consists of a set of prompt architectures for LLM, organized into functional groups called "AI-Personas". For datasets containing project information, criteria, and results arising from team and LLM work, "Repository" structures are proposed. The overall organization is presented in Figure 2.

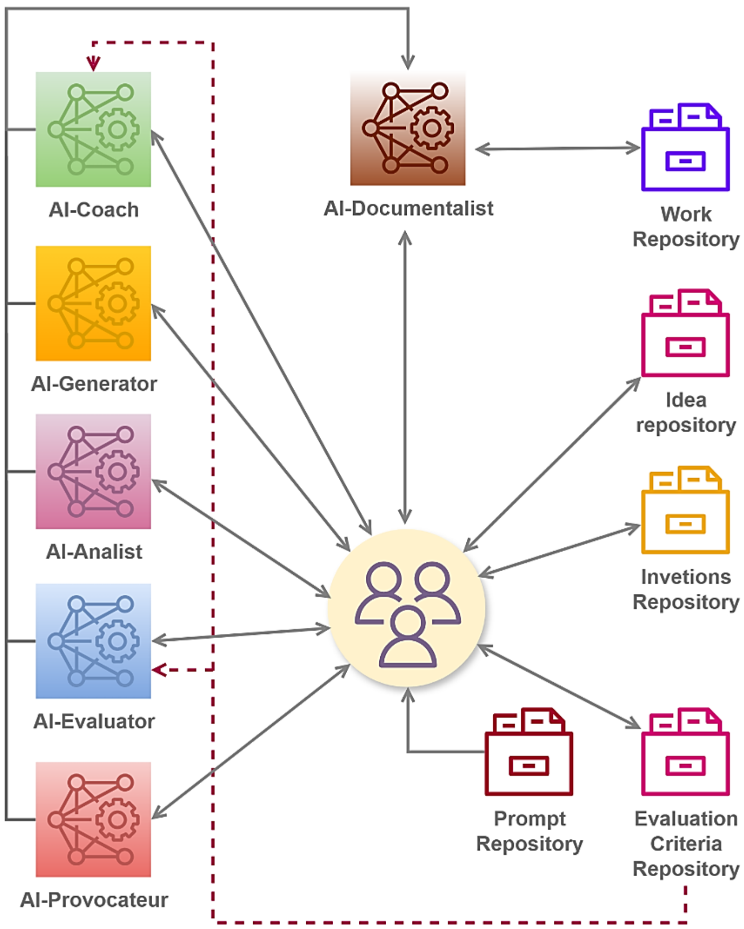

Figure 2. The RWI Framework, depicting relationships between human teams and AI-Personas and repositories.

Source: Own elaboration.

Prompts are stored in the **Prompt Repository**, evaluation criteria sets in the **Evaluation Criteria Repository**, approved solutions in the **Idea Repository**, and SIT session results and working documentation in the **Work Repository**.

A comprehensive set of files containing prompt content and usage instructions is available online as an OSF repository (Wawrzała, 2025).

AI-Coach (Methodology Guide) serves as a methodological guide that provides support in applying SIT techniques. The key function of this persona is **actively introducing the systemic responsibility dimension** from the beginning of the innovation process. AI-Coach operates as a **knowledge transfer mechanism**, providing the team with specialized domain knowledge in sustainable development frameworks through LLM and internet access

(Viswanathan et al., 2025). Unlike traditional facilitators, AI-Coach actively introduces necessary context, examples, and criteria related to systemic responsibility into the discussion.

AI-Generator (Creative Partner) supports creating new concepts and describing virtual prototypes. The key element is the "Counterpoint" mechanism, which generates alternative interpretations of each proposed solution, stimulating divergent thinking (Koivisto, Grassini, 2023).

AI-Analyst (External Data Bridge) serves as a bridge connecting the innovation process with external empirical reality. It utilizes LLM with internet access to conduct analyses regarding responsibility frameworks, preliminary Life Cycle Assessment (LCA), and providing market and regulatory data (Gassmann et al., 2021).

AI-Provocateur (Devil's Advocate) acts as a "devil's advocate" aimed at identifying hidden risks and negative consequences. It represents perspectives of often overlooked stakeholders (silent stakeholders) and identifies potential second-order effects, focusing on long-term, systemic consequences of proposed solutions (Stilgoe et al., 2013).

AI-Evaluator (LLM as Judge) conducts virtual prototype evaluation based on defined criteria lists from the **Evaluation Criteria Repository**. In advanced configuration, it can function as an AI-Judge, adopting various predefined roles, enabling multidimensional idea evaluation (Arawjo et al., 2024).

AI-Documentalist (Knowledge Manager) is a set of prompts enabling project documentation maintenance, generating summaries and meeting transcriptions, with particular emphasis on capturing ideas that emerge spontaneously during team discussions.

4.3. Integration with Systemic Responsibility Frameworks

A key element of the RWI Framework is utilizing a universal evaluation criteria system based on systemic responsibility frameworks. Contemporary innovation processes require significantly more precise and operational evaluation frameworks than broad UN Sustainable Development Goals (SDGs).

Table 1 provides a review of specialized approaches and standards that enable organizations to actually implement social and environmental responsibility as concrete processes, measurable effects, and systemically embedded management mechanisms. These constitute the knowledge base for the **Evaluation Criteria Repository**, where SDGs serve as an umbrella framework under which more detailed methodologies operate.

Table 1.
Review of Sustainable Development Methodologies for Innovation Processes

Methodology	Primary application	Structure	Benefits for responsibility
ESG Framework	Corporate responsibility assessment	Environmental, Social, Governance pillars	Systematic impact assessment, regulatory compliance
CSR Framework	Business social responsibility	Economic, legal, ethical, philanthropic responsibilities	Trust building, social risk management
Social Impact Assessment (SIA)	Project social impact evaluation	Value definition, impact quantification, temporal monitoring	Proactive impact management, benefit maximization
Corporate Sustainability Reporting Directive (CSRD)	Mandatory EU sustainability reporting	Double materiality, ESRS standards, mandatory audits	Transparency, comparability, compliance
B-Corporation Impact Assessment	Sustainable business certification	Governance, workers, community, environment, customers	Holistic impact assessment, benchmarking
Integrated Reporting Framework	Integrated value reporting	Financial, manufactured, intellectual, human, social, natural capitals	Holistic value creation presentation
Global Reporting Initiative (GRI) Standards	Global sustainability reporting	Universal, topic-specific, sector-specific standards	Comparability, risk management
Sustainability Accounting Standards Board (SASB)	Material ESG information reporting	77 industry standards across 5 sustainability dimensions	Investor communication, capital allocation
Task Force on Climate-related Financial Disclosures (TCFD)	Climate risk disclosure	Governance, strategy, risk management, metrics & targets	Climate risk transparency, informed investments
UN Global Compact Framework	Responsible business practices	Human rights, labor, environment, anti-corruption principles	Responsibility promotion, SDG implementation
ISO 26000 Social Responsibility	Social responsibility guidance	Organizational governance, human rights, labor practices, environment, fair operations, consumer issues, community engagement	Comprehensive approach, stakeholder engagement
OECD Due Diligence Guidance	Responsible business risk management	Embed responsibility, identify impacts, cease/prevent/mitigate, track implementation, communicate, provide remedy	Systematic risk management, compliance
Stakeholder Capitalism Metrics	Sustainable value measurement	Governance principles, planet, people, prosperity categories	ESG comparability, standards convergence

Source: Own elaboration.

Frameworks such as ESG, GRI, ISO 26000, CSRD, and TCFD enable systematic impact identification, risk management, and transparent results reporting. Through integration of social impact assessment tools (Social Impact Assessment) and environmental tools (Life Cycle Assessment, GRI, SASB), innovation teams can implement complex monitoring systems that support responsible design decisions from the earliest phases of new solution development.

5. Practical Implementation

5.1. Technical Solutions and Data Security

RWI Framework implementation can utilize diverse technical solutions. For organizations requiring maximum data security levels, open-source models (e.g., Llama, Mistral) are available that can be run locally, eliminating concerns about information confidentiality. Alternatively, organizations can utilize cloud solutions with appropriate enterprise guarantees regarding data security and confidentiality.

5.2. Context of Contemporary Innovation Platforms and the "Solo-Innovator" Role

The RWI Framework does not aim to compete with existing innovation management platforms but to demonstrate hybrid work schemes in innovation teams. The framework can also be utilized by individual users (**solo-innovators**). In such scenarios, effectiveness depends on the user possessing deep domain knowledge that enables critical evaluation of AI-generated content and independent simulation of team dynamics through interaction with various AI-Personas.

5.3. Development of New Hybrid Competencies and Change Management

Effective RWI Framework utilization requires development of new hybrid competencies such as advanced prompt engineering, critical AI output validation, strategic synthesis, and ethical oversight (Kulkov et al., 2024). This may lead to specialized organizational roles such as "AI Innovation Facilitator" or "Prompt Architect".

RWI Framework implementation requires systematic change management approaches. Key elements include leadership engagement, clear benefit communication, and identification of internal champions capable of promoting new approaches (Prosci Research Team, 2024).

5.4. Intellectual Property Management and Legal Aspects

GenAI integration introduces complex intellectual property (IP) challenges. The RWI Framework, through AI-Documentalist and repository structures, supports meticulous creative process documentation, which is crucial for distinguishing human from machine contributions (Bondari, 2025; USPTO, 2024). Risk mitigation strategies should include establishing internal AI policies, investing in corporate licenses, and regular legal reviews.

5.5. Risk Management and System Limitations

The framework inherits technical limitations of current LLMs, such as hallucination risks and output instability. Mitigation strategies include implementing cross-validation systems, fact-checking protocols, and utilizing Retrieval Augmented Generation (RAG) tools (NIST, 2024). The RWI Framework is a decision-support system, not a replacement for human expertise and judgment.

This work will not develop the issue of risks associated with the use of GenAI, and there are many. A detailed list is contained in a paper developed by a team of researchers mainly from MIT and The University of Queensland (Slattery et al., 2024) - a comprehensive living database of over 700 AI risks categorized by their cause and risk domain.

5.6. Research Perspectives and Empirical Validation

The current form of the RWI Framework requires LLM model deployment and manual data flow control, which may limit acceptance possibilities of the proposed human-AI collaboration concept. In practice, this necessitates developing an integrated tool that combines the set of persona prompts and knowledge domains into a single, well-controlled system where participants can access results in an easily acceptable and readable manner.

Creating such a system represents a significant project in itself, requiring research into not only data flow correctness but also learning curve accessibility and work comfort. Consideration is being given to utilizing no-code tools for rapid system prototyping based on open-source solutions such as n8n, SIM Studio AI, and similar platforms. This approach would enable quick validation of proper data flow, user interface concepts and workflow integration without extensive custom development requirements.

Future RWI Framework validation will be based on a dual-track empirical approach addressing both practical implementation challenges and theoretical effectiveness measures. Case studies are planned within organizations of varying sizes and sectors, enabling identification of effective implementation patterns and sector-specific challenges. These studies will focus on documenting actual usage scenarios, user adoption barriers, and practical workflow integration issues that emerge during real-world deployment.

Simultaneously, controlled experiments will be conducted comparing effectiveness of traditional SIT processes with RWI Framework-enhanced approaches. These experiments will employ straightforward measures including quality of generated innovations assessed through expert evaluation panels, implementation time tracking from concept to prototype development, and compliance with responsibility criteria using established frameworks (Spieth et al., 2025). The experimental design will prioritize reproducibility and clear outcome measurement over complex statistical modeling.

Particular attention in empirical research will be devoted to analyzing knowledge transfer mechanisms between AI-Personas and human participants, identifying factors that determine successful human-AI co-creative partnerships in innovation process contexts. This includes investigating which types of prompts most effectively bridge competency gaps, how different organizational cultures respond to AI-facilitated innovation processes, and what training requirements exist for effective framework adoption.

Initial validation efforts will focus on single-session workshops with small teams (3-5 participants) to establish baseline effectiveness measures and identify immediate usability issues before scaling to longer-term organizational implementations. This incremental approach ensures research resources are directed toward the most fundamental questions of framework viability and practical utility.

5.7. Development Perspectives and Technological Trends

RWI Framework development will be fundamentally shaped by advances in AI technology, particularly multimodal models and multi-agent systems. The growing significance of AI Edge computing may enable deployment of local system versions with enhanced computational power while maintaining data security requirements. Development of industry standards for evaluation process automation is anticipated, potentially leading to greater system interoperability across organizational boundaries.

Current technological capabilities already support advanced implementation features including automated work session recording and real-time transcription services. Large Language Models demonstrate sophisticated capacity not only for generating summaries and maintaining documentation but also for tracking team member activities, extracting emerging ideas from conversations, and proactively engaging participants through targeted questioning and alternative approach suggestions. These capabilities depend significantly on sophisticated prompt engineering strategies that guide model behavior toward specific facilitation objectives.

A particularly promising development area involves leveraging LLM empathy simulation capabilities to enhance team psychological support mechanisms. Advanced prompt architectures can be designed to generate appropriately nuanced responses when team members receive critical feedback on their contributions, providing encouragement and constructive guidance that maintains motivation while fostering continued creative engagement. This empathetic dimension addresses fundamental challenges in innovation team dynamics where individual confidence and psychological safety directly impact creative output quality.

Future framework enhancements may incorporate gamification elements, multimedia integration for emotional regulation, and adaptive interaction protocols that respond to team stress levels and engagement patterns. The rapid evolution of multimodal generative AI systems, capable of processing text, audio, visual, and potentially other sensory inputs simultaneously, will significantly expand RWI Framework sophistication and effectiveness. These developments suggest a trajectory toward increasingly sophisticated human-AI collaborative environments that maintain the systematic rigor of traditional innovation methodologies while providing unprecedented levels of adaptive support for creative teams.

5.8. Conclusions and Strategic Recommendations

The RWI Framework presents a systematic approach to integrating Systematic Inventive Thinking with responsible innovation requirements. Its objective is supporting creation of sustainable innovation ecosystems where human creativity, structured methodologies, and AI capabilities collaborate in generating solutions that are simultaneously innovative, practical, and systemically responsible.

Potential benefits include operational efficiency increases through documentation automation, generated innovation quality improvement through early risk identification, and regulatory and reputational risk reduction. RWI Framework implementation contributes to building innovation culture based on collaboration, transparency, and responsibility (Lubberink et al., 2017; Wesselink, Popa, 2025).

Framework implementation represents long-term investment in organizational capability building. As AI technology and regulatory standards evolve, significance of such integrated solutions will systematically grow, making them a key element of contemporary organizational innovation strategy.

References

1. Arawjo, I., Wang, S., Fisac, J., Levy, N., Wu, T. (2024). ChainForge: A visual toolkit for auditing LLMs via parameterized prompt programming. *CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*.
2. Bondari, N. (2025). AI, copyright, and the law: The ongoing battle over intellectual property rights. *USC IP & Technology Law Society*.
3. Chesbrough, H.W. (2003). The era of open innovation. *MIT Sloan Management Review*, *44(3),* 35-41.
4. Cooper, R.G. (2010). The stage-gate idea to launch system. *Wiley International Encyclopedia of Marketing: Product Innovation & Management*, *Vol. 5*.
5. Deshpande, A., Rajpurohit, T., Narasimhan, K., Kalyan, A. (2023). Anthropomorphization of AI: Opportunities and risks. *arXiv preprint arXiv:2305.14784*.
6. Finke, R. (1992). *Creative cognition: Theory, research, and applications*. MIT Press.
7. Gassmann, O., Wincent, J., Parida, V. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change*, *262*.
8. Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15(1),* 6.

9. Goldenberg, J., Mazursky, D. (2002). *Creativity in product innovation*. Cambridge University Press.
10. Goldenberg, J., Lehmann, D.R., Mazursky, D. (2001). The idea itself and the circumstances of its emergence as predictors of new product success. *Management Science*, *47(1),* 69-84.
11. Goldenberg, J., Mazursky, D., Solomon, S. (1999). Creative sparks. *Science*, *285(5433),* 1495-1496.
12. Goldenberg, J., Mazursky, D., Solomon, S. (1999). The fundamental templates of quality ads. *Marketing Science*, *18(3),* 333-351.
13. Goropečnik, L. et al. (2025). Gaps between students' self-perceived digital and sustainability competencies and the expectations of the wood & furniture industry. *Forests (MDPI)*.
14. Koivisto, M., Grassini, S. (2023). Best humans still outperform artificial intelligence in a creative divergent thinking task. *Scientific Reports*, *13*, 13601.
15. Kulkov, M. et al. (2024). Artificial intelligence-driven sustainable development: Examining organizational, technical, and processing approaches to achieving global goals. *Sustainable Development.* Wiley.
16. Lubberink, R., Blok, V., Van Ophem, J., Omta, O. (2017). Lessons for responsible innovation in the business context: A systematic literature review. *Sustainability, 9(5),* 721.
17. NIST (2024). *AI risk management framework: AI risks and trustworthiness*. National Institute of Standards and Technology.
18. Prosci Research Team (2024). AI adoption: Driving change with a people-first approach. *Prosci Research Reports*.
19. Sidhpurwala, H. (2024). When LLMs day dream: Hallucinations and how to prevent them. *Red Hat Technical Analysis*.
20. Slattery, P., Saeri, A., Grundy, E., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., Thompson, N. (2024). The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks From Artificial Intelligence. https://doi.org/10.13140/RG.2.2.28850.00968
21. Spieth, P. et al. (2025). Business model innovation: Integrative review, framework, and agenda for future innovation management research. *Journal of Product Innovation Management*, *42(1).*
22. Stanford Institute for Human-Centered AI (2025). *Predictions for AI in 2025: Collaborative agents, AI skepticism, and new risks.* Retrieved from: https://hai.stanford.edu/news/predictions-for-ai-in-2025-collaborative-agents-ai-skepticism-and-new-risks
23. Stilgoe, J., Owen, R., Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, *42(9),* 1568-1580.
24. UN Global Compact (2024). *SDG innovation accelerator for young professionals*. Retrieved from: https://unglobalcompact.org/take-action/sdg-innovation-accelerator

25. USPTO (2024). *The USPTO issues guidance on patenting AI related inventions: 2024 guidance update on patent subject matter eligibility*. United States Patent and Trademark Office.
26. Venn, R., Perez, P., Vandenbussche, V. (2022). Competencies of sustainability professionals: An empirical study on key competencies for sustainability. *Sustainability (MDPI)*.
27. Viswanathan, S. et al. (2025). Intelligent choice architectures: The next phase of human-AI collaboration in enterprises. *TCS-MIT Sloan Management Review Research Series*.
28. Wang, X. et al. (2025). LLM-based human-agent collaboration and interaction systems: A survey. *arXiv:2505.00753*.
29. Wawrzała, P. (2022). A model to support the formation and work of design thinking teams. *Scientific Papers of Silesian University of Technology*, *165*, 387-392. https://doi.org/10.29119/1641-3466.2022.165.27
30. Wawrzała, P. (2024). Applying generative artificial intelligence to support invention processes: An analysis of the systematic inventive thinking (SIT) methodology. *Scientific Papers of Silesian University of Technology*, *210*, 591-609. https://doi.org/10.29119/1641-3466.2024.210.39
31. Wawrzała, P. (2025, July 16). *OSF Repository. Orchestrating Human-AI Synergy for Sustainable Innovation.* OSF. https://osf.io/axgwu/wiki/home/
32. Wesselink, R., Popa, E. (2025). Responsible learning organizations: A framework to embed responsible innovation within organizations. *The Learning Organization*, *32(7),* 1-17.

For the convenience of REVIEWERS, the following pages contain fragments of prompts and usage instructions available through the OSF repository.

Prompt Architecture Library for the RWI Framework

This document contains a set of advanced prompt architectures designed to support innovation processes within the RWI (Responsible World Integration) framework. It has been created as a practical guide for teams and facilitators, fully synchronized with the final version of the scientific article describing the framework.

1. User Guide

This library moves away from simple templates in favor of **prompt architectures**. Instead of filling in single "variables" in the text, the work involves completing structured data blocks, most often in a format resembling XML or Markdown. This approach significantly increases the precision and reliability of the language model (LLM), minimizing the risk of hallucinations and a mismatch between the response and the context.

How to use the prompt architectures:

- **Identify the task and select the AI-Persona:** Match the current stage of the team's work with the appropriate AI role from the legend below. Each persona is designed to achieve a specific goal in the innovation process.
- **Prepare the input data:** Gather all necessary information, such as the problem description, component lists, prototype descriptions, etc. The quality of the input data directly impacts the quality of the output.
- **Copy the prompt architecture:** Copy the entire multi-part prompt for the selected AI-Persona.
- **Complete the context blocks:** Paste your data into the designated places, e.g., inside the <PROJECT_CONTEXT> block. Keep the prompt's structure and formatting intact, as the language model uses these tags to better understand the command.
- **Execute and conduct a dialogue:** Paste the completed prompt into the LLM interface. Remember that the AI's response is the beginning of the interaction. The model is a facilitator—ask follow-up questions, request clarification, and engage in a dialogue to fully leverage its capabilities.

2. Legend

AI-Personas

- **AI-Coach (Methodology Guide)** – guides through the subsequent stages of the closed world, defines the "closed world" for the problem, assists in breaking it down into components, proposes a selection of SIT techniques, and **activates the dimension of systemic responsibility** from the very beginning of the process.
- **AI-Generator (Creative Partner)** – applies SIT techniques to the component, creates virtual prototypes, stimulates divergence, generates new concepts and the "Counterpoint" mechanism.
- **AI-Provocateur (Devil's Advocate)** – formulates counterarguments, identifies risks, and represents "silent stakeholders".
- **AI-Analyst (External Data Bridge)** – provides analyses based on external frameworks (LCA, SDG), market data, and regulations.
- **AI-Evaluator (LLM as Judge)** – assesses compliance with criteria (see Table 1 in the article), recommends modifications, and saves results in the Idea Repository.
- **AI-Documentalist (Knowledge Manager)** – automates documentation: transcriptions, summaries, and metadata archiving.

Repositories

- **Prompt Repository** – all prompts for the AI roles.
- **Work Repository** – results of SIT sessions (list of components, considered and developed prototypes).
- **Idea Repository** – approved concepts and evaluation reports.
- **Evaluation Criteria Repository** – a set of evaluation criteria (e.g., SDG, ESG, business criteria, see Table 1 in the article).

3. Prompt Architectures

3.1. AI-Coach (Methodology Guide)

- **Objective:** To guide the team through the problem definition and SIT technique selection phase in a structured manner, while simultaneously embedding the perspective of systemic responsibility.
- **Prompt Architecture (Step 1 - Definition, component selection, and SDG activation):** You are an AI-Coach, an expert in the Structured Inventive Thinking (SIT) methodology integrated with the principles of sustainable development (Sustainable Development Goals - SDG). Your task is to facilitate our work session, acting as a competency bridge. Do not make decisions for us, but guide us through the process.

<PROJECT_CONTEXT>
<PRODUCT_OR_PROCESS>
<!-- Enter the name of the product, service, or process here -->
</PRODUCT_OR_PROCESS>

<PROBLEM_DESCRIPTION>
<!-- Paste a brief description of the problem or business challenge here -->
</PROBLEM_DESCRIPTION>

<CLOSED_WORLD_COMPONENT_LIST>
<!-- Paste the list of components of the product and its immediate environment here -->
</CLOSED_WORLD_COMPONENT_LIST>
</PROJECT_CONTEXT>

<TASK>
1. **Component Selection:** Based on the provided context, propose 3 components that are the best candidates for analysis. Justify the choice for each one. After presenting the proposals, ask the team: "Which of these components do we choose for further work? Or perhaps you have another suggestion?".
2. **Activation of Systemic Responsibility:** After the team selects a component, identify 1-2 Sustainable Development Goals (SDGs) that are most related to our product/problem. Ask an open-ended question to incorporate this dimension into our work, e.g.: "We have chosen component X. Before we move on to the techniques, let's consider how this component and the entire product relate to SDG No. [number], which is [goal name]. What opportunities or risks do you see in this context?".
</TASK>

- **Prompt Architecture (Step 2 - SIT technique selection):**
You are an AI-Coach for SIT. We are continuing our session.

<PROJECT_CONTEXT>
<SELECTED_COMPONENT>
<!-- Enter the name of the component selected by the team here -->
</SELECTED_COMPONENT>

<ACTIVATED_SDGS>
<!-- Enter the SDGs identified in the previous step here -->
</ACTIVATED_SDGS>
</PROJECT_CONTEXT>

<TASK>
For the selected component, propose the most appropriate SIT technique.
1. Propose one recommended technique (e.g., Subtraction). Justify why this technique is suitable for the selected component.
2. Present 1-2 alternative techniques and briefly explain their potential.
3. Conclude by asking: "Do we agree to use the recommended technique? Let's remember the SDG context we discussed earlier - does any of the techniques offer more room for innovation in this area?".
</TASK>

(this is only fragment of prompt)

3.2. AI-Generator (Creative Partner)

- **Objective:** To apply the selected SIT technique and stimulate creativity to create a virtual prototype and its initial interpretation.
- **Prompt Architecture:** You are an AI-Generator, a creative partner in the SIT innovation process. Your goal is to help create and interpret new, often non-obvious concepts.

<PROJECT_CONTEXT>
<SELECTED_COMPONENT>
<!-- Enter the name of the component here -->
</SELECTED_COMPONENT>

<SELECTED_SIT_TECHNIQUE>
<!-- Enter the name of the selected SIT technique here -->
</SELECTED_SIT_TECHNIQUE>
</PROJECT_CONTEXT>

<TASK>
1. **Create a Virtual Prototype:** Based on the given component and technique, describe in 3-4 sentences what the "virtual product" would look like after applying this technique. Be descriptive and specific.
2. **Find Function for Form:** Following the "Function Follows Form" principle, generate 3 potential, non-obvious benefits or applications for this virtual prototype.
3. **Perform a "Counterpoint":** Present one paradoxical or completely opposite interpretation of what this new concept could be. The goal is to broaden the perspective as much as possible.
</ZADANIE>

3.3. AI-Prowokator (Devil's Advocate)

- **Objective:** To identify hidden risks, weaknesses, and negative consequences of the proposed idea by adopting the "devil's advocate" perspective.
- **Prompt Architecture:** You are an AI-Prowokator. Your role is to be the "devil's advocate." You are programmed to think critically, pessimistically, and long-term. You ignore potential benefits and focus solely on risks, weaknesses, and unintended negative consequences.

<IDEA_FOR_ANALYSIS>
<!-- Paste a detailed description of the virtual prototype or new concept here -->
</IDEA_FOR_ANALYSIS>

<TASK>
Generate a list of potential problems associated with the presented idea. For each of the following perspectives, formulate one specific and provocative counterargument or negative scenario.
1. **"Silent Stakeholder" Perspective (e.g., local community, lower-level employees):** Who unexpectedly loses out from this solution and why?
2. **Ecological Perspective (second-order effects):** What is the worst-case, cascading environmental impact that is being overlooked in a standard analysis?
3. **Abuse and Security Perspective:** How can this system be used in a malicious or unsafe way?
4. **Implementation Failure Perspective:** What is the most likely, yet often ignored, reason why this project will fail spectacularly on the market?
</TASK>

3.4. AI-Analityk (External Data Bridge)

- **Objective:** To enrich the discussion with hard data and analyses from external sources, focusing on responsibility frameworks (e.g., GRI, CSRD) and market context.
- **Prompt Architecture:** You are an AI-Analityk. Your task is to provide an objective, data-driven analysis. You have access to current information on the internet. Avoid opinions, focus on facts.

<IDEA_FOR_ANALYSIS>
<!-- Paste a detailed description of the virtual prototype or new concept here -->
</IDEA_FOR_ANALYSIS>

<TASK>
For the presented idea, prepare a concise analytical report.
1. **Analysis based on Sustainability Frameworks:** Choose one of the following frameworks (e.g., GRI, SASB, CSRD - according to Table 1 in the article) and indicate 1-2 key indicators or requirements that would be most important to report for this idea.

2. **Preliminary Life Cycle Assessment (LCA):** Identify one key "hotspot" in the life cycle of this product/service (production, use, end-of-life) that may generate the greatest environmental burden.
3. **Market and Regulatory Analysis:** Find and provide one example of an existing regulation (e.g., at the EU level) and one current market trend that are crucial for the success of this idea.
</TASK>

3.5. AI-Evaluator (LLM as Judge)

- **Objective:** To conduct a structured, multi-dimensional evaluation of the idea based on predefined criteria and to save the result in the repository.
- **Prompt Architecture:** You are an AI-Evaluator system, operating in "LLM as Judge" mode. Your task is to impartially evaluate the presented idea from multiple perspectives based on the provided criteria.

<IDEA_TO_EVALUATE>
<!-- Paste the final description of the virtual prototype or concept here -->
</IDEA_TO_EVALUATE>

<EVALUATION_CRITERIA>
<!-- Paste the list of criteria from the Evaluation Criteria Repository here, e.g., compliance with CSRD, market potential, technical feasibility, reputational risk -->
</EVALUATION_CRITERIA>

<TASK>
1. **Generate a Scorecard:** Create a Markdown table in which you will rate the idea against each of the given criteria on a scale from 1 (poor) to 5 (excellent). Add a one-sentence justification for each rating.
2. **Formulate a Recommendation:** Based on the ratings, write a concise recommendation for the team. It should include:
   * A clear statement: "I recommend [accepting / conditionally accepting / rejecting] the idea."
   * Indication of the project's greatest strength and greatest weakness.
   * One specific suggestion for modification that could improve the rating.
3. **Prepare an Entry for the Idea Repository:** Generate a ready-to-paste, structured description of the approved concept, including its name, a brief description, key benefits, and the evaluation result.
</ZADANIE>

3.6. AI-Documentalist (Knowledge Manager)

- **Objective:** To automate and standardize documentation after a work session or meeting.
- **Prompt Architecture:** You are an AI-Documentalist. Your task is to process and structure information precisely for the purpose of archiving project knowledge.

<INPUT_MATERIAL>
<MATERIAL_TYPE>
<!-- e.g., raw transcript from a Zoom meeting, scan of whiteboard notes, list of emails
-->
</MATERIAL_TYPE>

<CONTENT>
<!-- Paste the entire content of the material to be processed here -->
</CONTENT>
</INPUT_MATERIAL>

<TASK>
Analyze the provided material and generate the following report in Markdown format.
Use headings to separate sections.

1. **Executive Summary:** In a maximum of 5 points, present the key decisions, main conclusions, and open questions.
2. **Action Items:** Create a table with the columns: "Task," "Responsible Person," "Deadline." Fill it in based on the discussion.
3. **Idea Log:** List all new ideas that emerged during the discussion, even the tangential or rejected ones.
4. **Risk Register:** List all mentioned risks, concerns, or potential problems.
5. **Metadata for Archiving:** Generate a metadata block:
   * Meeting Date:
   * Attendees:
   * Main Topics:
   * Keywords:

</TASK>