KM4ESG: BPMN and AI-powered knowledge management platform for ESG analysis and reporting

Damaris-Naomi Dolha[1,†], Cristina-Claudia Osman[1,†,*] and Andrei Chiș[1,†]

[1]*OMILAB@UBB-FSEGA & Business Informatics Research Center, Faculty of Economics and Business Administration, Babeș-Bolyai University, 58-60, T. Mihali Street, Cluj-Napoca, Romania*

Abstract

This paper outlines a project vision and proposal submitted to a Romanian national grant competition on sustainability. The objective is to engineer an ESG (Environmental, Social and Governance) Knowledge Management capability that leverages multi-modal knowledge representation – BPMN diagrams enriched by an ESG-specific semantic layer, converted to knowledge graphs and exposed to LLM services through a Graph RAG architecture. A proof-of-concept of technological readiness level 4 is planned to be developed, to demonstrate the streamlining of the different knowledge representation modes and to enable evaluative experimentation. The research is framed as a Design Science research project due to its artifact-building nature and practical problem orientation.

The problem pertains to a need for a Knowledge Management capability identified through several regional workshops discussing ESG requirements and challenges for local small and medium enterprises. Data analytics tools and reporting templates are typically invoked in such contexts, but most are repurposed legacy tools, lacking a granular mapping on enterprise architecture layers or business operations. Such mappings are relevant because ESG injects new architecture elements (new ESG-oriented roles, processes, documents), new semantics (new task types, event types, flows), new attributes (e.g. carbon footprint) which are all in the scope of enterprise modeling – if not at architecture level, at the very least at business process management level. LLM agents must become aware of such semantics to tailor their content or recommendation generation for enterprise knowledge already available in BPMN and other types of enterprise models. However, legacy enterprise modeling methods are semantically insufficient to capture cross-cutting ESG concerns, so a knowledge engineering effort is also included in this project vision, to extend BPMN with the domain-specificity of ESG and to achieve a "meet-in-the-middle" point between ESG accounting requirements and Generative AI content.

Keywords

ESG, Domain-Specific Modeling, Business Process Management, Knowledge Graphs, Large Language Models

1. Introduction

Participation in ESG-focused workshops organized by industry clusters in the Cluj-Napoca, Romania [1] served as a starting point for this project proposal, developed as a 20 months work plan for a competition launched in 2025 by the Romanian Academy of Romanian Scientists[1].

The regional workshops highlighted the growing demand for advanced ESG management tools that go beyond the scope of data collection and reporting, towards managing the intricate dependencies involved in ESG data provisioning through cross-cutting concerns across layers of enterprise and business process architectures.

[1] https://www.aosr.ro/en/

ESG reporting (also known as sustainability reporting) is becoming a mandatory requirement for companies. The reporting effort is complex, requiring the orchestration of data sources, new work procedures, rules and multi-criterial decision-making. ESG reporting is defined at international level (through standards and regulations such as International Accounting Standards IAS/IFRS, Directive 2014/95/EU (NFRD), Directive 2464/2022 – CSRD, the Paris Agreement, B Corp Certification, etc.) and at national level it involves the incorporation of European ESG directives into national laws. In Romania, this adaptation includes the Ministry of Public Finance Order no. 85 (January 12, 2024) [2], mandating sustainability reporting for entities with over 500 employees.

This national framework triggered many regional workshops with local IT service providers raising both awareness and requirements. In our previous reports [1], we have contrasted the dominant data-centric products (analytics reporting tools repurposed for ESG) against the generally unavailable Knowledge Management capabilities required to support ESG-focused roles and work systems.

In the traditional practice of Business Process Management (BPM), organizations consider factors like time, cost, efficiency and quality. However, these factors do not directly ensure sustainability – business processes must be augmented with dedicated task types, execution roles or data object taxonomies to serve ESG goals that are additional to the main process goal. A new dimension has been added to BPM, leading to the emergence of Green Business Process Management (Green BPM) [3] which highlights the environmental impact of business processes. This allows for specific improvements in the E&S dimensions of ESG, recognizing that the G dimension is generally well-managed through existing BPM methodologies.

For technological platforms, Generative AI can play a key role in helping organizations to design and execute a structured, content-driven ESG roadmap enabling better decision-making, enhanced reporting, and alignment with sustainability goals [4]. Therefore, our project vision addresses *the need to integrate ESG capabilities into business process descriptions*, while at the same time exposing such enriched descriptions to Generative AI assistance - to better support process analysis, reporting and generally process-centric Knowledge Management. There seems to be a practitioner experience gap between ESG intentions and measurable outcomes due to fragmented knowledge, lack of traceability, and the complexity of ESG metrics. Our project vision seeks to bridge that gap by proposing the development of a hybrid AI platform that will leverage BPMN as a process description standard, ESG as a semantic enrichment layer (on metamodel level) and knowledge graphs as mediators to inform Large Language Model (LLM) services with process-centric ESG design decisions and semantic annotations.

The paper is structured as follows: this introductory unit is followed by related work comments focusing on the BPM lifecycle integrated with LLMs. The problem statement and research methodology are depicted in Sections 3 and 4, followed by an outline of preliminary efforts in combining BPM and a hybrid LLM-KG approach. The research vision is presented in Section 6, making explicit the relevance to CAiSE in Section 7, followed by conclusions.

2. Related works

ESG helps to assess how an organization manages risks and opportunities created by cross-cutting factors along three dimensions. The **Environmental** dimension measures a company's impact on the natural ecosystem (carbon emissions, efficient resource utilization in production processes, pollution, waste, eco-friendly efforts and sustainable services). This dimension is influenced by the other two – **Social**, referring to the company's social relationships, both external (along supply chains or activities with local communities) and internal (with employees); and **Governance**, which concerns transparency and responsibility within company's operations.

Given the complexity of semantic intricacies between ESG factors, organizations face significant challenges in integrating and tracing ESG considerations into business processes. GenAI solutions are valuable in facilitating ESG data measurement, e.g. in areas like emissions tracking or assessing social aspects like gender diversity [4], however *we advocate the use of GenAI beyond data analytics scenarios, since it can also handle conceptual (process) analysis* [5].

The use of GenAI as a tool to assist business analysts and automation is demonstrated in [6]. Starting from diagrammatic visualizations (such as ER and BPMN diagrams), GenAI tools can provide support throughout the software development lifecycle (SDLC). Our proposal does not focus on SDLC, but rather on the BPM lifecycle [7]. The potential of integrating LLMs into each stage of the BPM lifecycle was suggested by the agenda of Vidgof et al. [8]. Going beyond that call to action, this research also employs metamodeling to enrich business process modeling with ESG aspects currently not covered by any graphical modeling standard — consequently, a key contribution of this proposal will be the extension of BPMN with distinctive ESG constructs. Early stage propositions are based on the Bee-Up modeling tool[2] as reported in [1].

Figure 1 describes the traditional BPM lifecycle and maps it to potential contributions of LLMs in the *Process identification* phase, LLMs can analyze regulations and policies with the goal of identifying business processes into which ESG factors can be integrated. Grohs et al. [9] use LLMs in process identification and discovery from textual descriptions, their evaluation indicating promising performance. Including LLMs in the *Process discovery* phase helps on mapping ESG initiatives within existing business processes; at this phase, LLMs can be viewed as Process Mining tools [10] focused on discovering processes based on historical data.
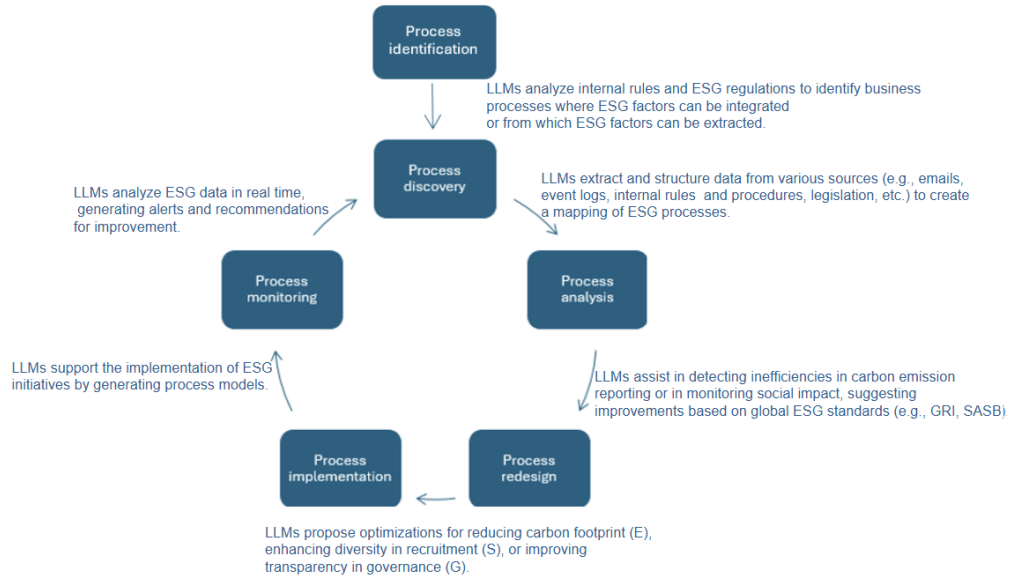
Figure 1: BPM lifecycle [7] integrated with LLMs [8] in the context of ESG reporting

Process analysis with the help of LLM can help with identification of ESG risks or greenwashing [11]. *Process redesign* with LLM support includes ESG policy recommendations (for example, reducing energy and water consumption in certain processes). *Process implementation* integrated with LLM may include an automatic generation of ESG reporting or guiding employees in executing manual work procedures through LLM-generated recommendations. *Process monitoring* with LLM provides suggestions based on ESG data analysis, monitoring compliance with current regulations, as well as evaluating ESG impact or generating predictions.

Another converging research stream investigates how GenAI models can parse BPMN diagrams either in graphical form or serialized through various export formats [5], which leads to our suggestion of employing semantic graphs as mediator and interchange format between a BPMN modeling tool and an LLM service. By extending existing standards like BPMN with ESG concepts, we also face the challenge of how well metamodel augmentations are received by LLMs since no training corpus is available for such narrow scope domain-specific modeling languages (DSMLs). We aim to experiment on this challenge by informing the LLM with various RDF-based process

---

vocabularies extended with ESG-specific semantics – e.g. new taxonomies of tasks, events, flows, data objects.

3. Problem statement and envisioned solution summary

ESG evaluations must be primarily facilitated internally, providing organizations with tools for assessing and designing their own ESG policies tightly coupled to business operations. This motivates the objective of this research – the development of an innovative **AI and BPMN-based platform for Knowledge Management for ESG reporting**. Integration with GenAI will answer gaps such as what Minkkinen et al. highlight as a lack of responsible AI integration in ESG investment analysis [12].

The idea of engineering a platform that leverages BPMN and GenAI for ESG Knowledge Management relates to the current state of widely adopted ESG tools [1]. Morningstar Sustainalytics [13] and Refinitiv [14] calculate ESG scores based on data provided by companies and compare them with other companies within the same industry sector. These tools rely on standards and frameworks such as the Global Reporting Initiative [15], the Sustainability Accounting Standards Board [16] and the World Economic Forum [17]. However, their ESG analysis is based on aggregating quantitative attributes from various sources, without analyzing or ensuring the traceability of those data sources to organizational workflows or supply chain dependencies, with no means to decompose and granularly map the ESG concerns. Granular task-resource-data mappings can be found in the general BPM practice for simulation purposes and we consider how this can be repurposed for ESG accounting – not only in terms of data attributes, but also in terms semantic enrichments and process dependencies on ESG elements.

Therefore, our project vision aims to repurpose and extend BPMN to complement the data-centric quantitative ESG practices with an interpretative platform based on hybrid AI interpreting ESG-enriched business process models. Identifying and managing the interplay between ESG pillars, business processes and enterprise systems can lead to the Knowledge Management capability targeted by this project proposal:

- The ESG-process dependencies are chains of relationships for which As-Is analysis and To-Be prescriptive process redesigns are required – therefore, we use the BPMN standard as a starting conceptual structure, to be extended with new taxonomies of tasks, events, decisions, roles, information flows and dependencies that capture the ESG perspective;
- GenAI tools are increasingly recognized as advisory services (recommendations) or for analysis (of private content). Hybridizing LLMs and knowledge graphs to optimize an LLM for an organization's private content [18] can be further enhanced with diagrams inherited from traditional BPM practices [5,19,20].

By combining these two technological facets, this project proposal adopts an updated form of "Model-driven Engineering" (MDE) - promoted as "Semantics-driven Engineering" [21] due to the focus on knowledge streamlining between diverse knowledge processing environments. Instead of focusing on diagrams to generate software components or automate workflows, we extend the BPMN standard to inform GenAI assistants with metamodel-enriched process semantics that incorporate ESG concerns in the process description vocabulary.

## 4. Project plan

4.1. Objectives

The objective of the project proposal is to develop an intelligent platform based on AI and BPMN as support in Knowledge Management on the organization and tracing of ESG-related activities. The targeted requirements are to facilitate the design, documentation, analysis and optimization of

business processes from the perspective of ESG concerns. The proposed innovative platform will consist of the following components:

- a graphical language (DSML) and the corresponding modeling tool that extend the BPMN standard [22] with ESG-specific concepts, dependencies and taxonomies,
- automated mechanisms for analysis, reporting and decision support through the interoperability of the extended BPMN+ESG diagrams with GenAI services (OpenAI services to be used for a demonstrator),
- interoperability mechanisms based on RDF graphs as a mediator between the modeling tool and the GenAI services,
- an integrated operating procedure to support a managerial capability that can be evaluated using Knowledge Management maturity standards such as APQC's maturity model [23].

4.2. Methodology

The work methodology is tailored for an artifact-building engineering effort that involves both "method engineering" in the sense of [24] and software development. The appropriate methodology for this type of scientific research is Design Science Research (DSR) [25]. Considering the specific task of extending an existing modeling language, the AMME methodology (Agile Modeling Method Engineering) [26] is also needed and will rely on the ADOxx metamodeling platform. Therefore, the project methodology reflects the hybridization between DSR and modeling method engineering, as suggested in Fig. 2.
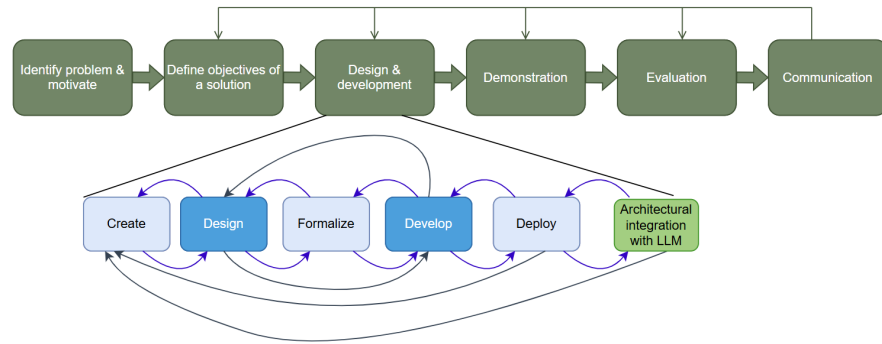
Figure 2: Hybrid approach: DSR and AMME (adapted after [25] and [26])

4.3. Work plan

The DSR methodology structures the project's approach into well-defined stages that also inform the work plan, which is inherently iterative, allowing for the gradual expansion of the developed artifact based on intermediate feedback. This requires two engineering-evaluation cycle iterations (WP3 and WP4) as the core of the work plan detailed below:

- **WP1.** Survey literature and practitioner feedback
- **WP2.** Consolidate solution requirements
- **WP3.1.** Design and development of the BPMN-ESG modeling tool
- **WP3.2.** Design and development of the model-driven AI-based KM platform
- **WP3.3.** Demonstration and evaluation of the integrated prototype
- **Milestone 1 (Month 10)** – Evaluated prototype on Technological Readiness Level 3
- **WP4.1.** Design and development of the BPMN-ESG modeling tool
- **WP4.2.** Design and development of the model-driven AI-based KM platform
- **WP4.3.** Demonstration and evaluation of the integrated prototype
- **Milestone 2 (Month 20)** – Working prototype on Technological Readiness Level 4

- **WP5.** On-going dissemination for Months 1-20.

5. Preliminary efforts

The formulation of this project proposal was preceded by several converging activities: As reported in [1] ESG workshops hosted by local industry clusters in our region highlighted the need for ESG Knowledge Management and data traceability capabilities. This motivated a survey on ESG tooling emphasizing the limitations of data-driven ESG tools and the need for ESG traceability [27]. On the technological level, architectural configurations that allow the hybridization of LLMs and knowledge graphs have been identified in the GraphRAG family of patterns[3], building on our previous work of deriving such knowledge structures and RDF from diagrammatic designs [5,6,19,20]. Our earlier results comparing ChatGPT's interpretation of BPMN models in RDF and XML formats, suggested that RDF graphs are superior for navigating complex process dependencies and work better in an open world assumption where additional semantic contexts can agilely expand the procedural knowledge [5]. Another work stream investigated the representation of business process models as vectors [20].

6. Artifact components and envisioned architecture

Fig. 3 illustrates the architectural vision of the platform to be developed, including a toy example for a bottle recycling scenario modelled on the left side. The ESG-specific modeling tool and DSML developed in the ADOxx metamodeling platform[4] will enable the representation of ESG-enriched processes – i.e. BPMN processes with ESG task taxonomies, event taxonomies, pool (role) taxonomies, flow taxonomies (e.g. waste flows) and associated data objects. The resulting models are to be converted to RDF graphs (extending the mechanism involved in [5]) and exposed to LLM services through different prompting strategies based on the Teller taxonomy [4]. Using a Knowledge Graph as a mediator, the platform transfers diagrammatic process knowledge through a structured form that LLMs can interpret with better precision. Experimentation will employ the RAGAs metrics[5] to assess the accuracy of LLM responses informed by ESG-specific BPMN models exchanged as RDF graphs.

7. Relevance to CAiSE

The KM4ESG project aligns to several topics of interest for the CAiSE community:

- *Artificial Intelligence including generative AI and Machine Learning:* i.e. the application of AI techniques to enhance information systems;
- *Knowledge Graphs*: the use of Knowledge Graph as mediator to improve LLM interaction highlights the conference's interest in decision models and AI integration;
- *Business process modeling, analysis and improvement:* focusing on traditional BPMN and later incorporating ESG-specific elements (extending BPMN for ESG integration) ensures that business process descriptions are the core procedural knowledge for ESG concerns;
- *Sustainability and social responsibility management:* the integration of ESG into business process supports decision-making aligned with sustainability goals.
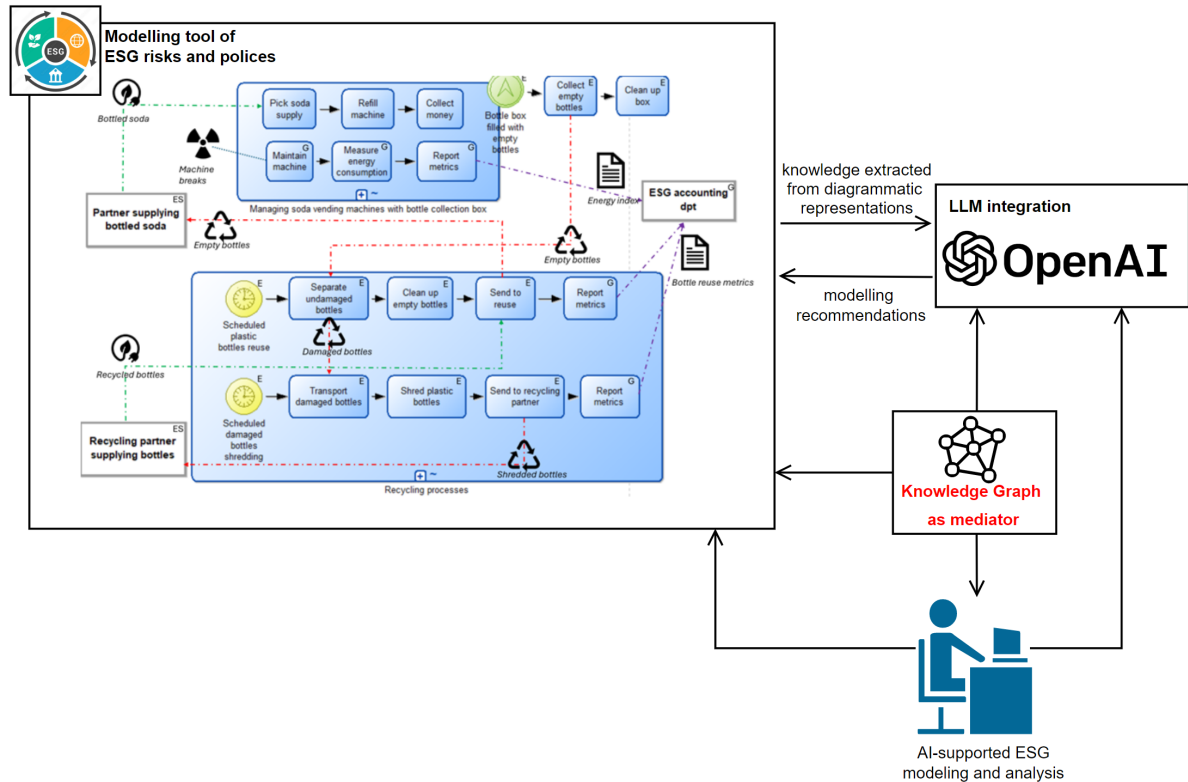
Additionally, the proposal connects directly to CAiSE's call for the development and evaluation of artifacts that advance both theory and practice in information systems engineering.

---

[3] https://www.ontotext.com/knowledgehub/fundamentals/what-is-graph-rag/
[4] https://adoxx.org/
[5] https://docs.ragas.io/en/stable/

Figure 3: Architectural vision

8. Conclusions

AI can play a crucial role in ESG reporting, offering significant benefits for both companies and stakeholders [29,30]. GenAI automation not only reduces the volume of manual labor but also ensures the accuracy of reporting in compliance with current regulations [31].

This project vision proposes the development of a model-driven, process-centric, platform for Knowledge Management and an updated view of traditional knowledge flows (e.g. Nonaka's conversion cycle [32]). The envisioned contributions target two areas: enriching domain-specific modeling by integrating ESG considerations into BPMN and improving sustainability management with the help of hybrid AI informed through a model-driven approach.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] Osman, C.C., Ghiran, A.M., Buchmann, R.A., Towards a Knowledge Management Capability for ESG Accounting with the Help of Enterprise Modeling and Knowledge Graphs. in: Companion Proceedings of the 17th IFIP WG 8.1 Working Conference on the Practice of Enterprise Modeling Forum, M4S, FACETE, AEM, Tools and Demos co-located with PoEM 2024, CEUR-WS, Aachen, 2024.

[2] Ministry of Finance, Order no. 85 of January 12, 2024, on regulating aspects related to sustainability reporting, 2025. URL: https://legislatie.just.ro/Public/DetaliiDocument/278502.

[3] Seidel, S., Recker, J., vom Brocke, J., Green Business Process Management. In: Green Business Process Management: Towards the Sustainable Enterprise, Springer Berlin, Heidelberg, pp. 3–13, 2012, doi: https://doi.org/10.1007/978-3-642-27488-6.

[4] EY, How Generative AI can build an organization's ESG roadmap, 2025. URL: https://www.ey.com/en_in/insights/ai/how-generative-ai-can-build-an-organization-s-esg-roadmap/.

[5] Dolha, D.N., Buchmann, R.A., Generative AI for BPMN Process Analysis: Experiments with Multi-Modal Process Representations. In: Perspectives in Business Informatics Research. BIR 2024. Lecture Notes in Business Information Processing, vol 529. Springer, Cham, 2024, pp. 19-35. doi: https://doi.org/10.1007/978-3-031-71333-0_2.

[6] Niculescu, V., Chisăliță-Crețu, M.C., Osman, C.C., Sterca, A., Model-Driven Development Using LLMs: The Case of ChatGPT. In: Proceedong of the 20th International Conference Evaluation of Novel Approaches to Software Engineering, Science and Technology Publications, Lda, Setúbal, Science and Technology Publications, Lda, Setúbal, 2025, pp. 328-339. doi: 10.5220/0013484400003928.

[7] Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A., Fundamentals of Business Process Management, Second edition. Springer, Heidelberg, 2018. doi: https://doi.org/10.1007/978-3-662-56509-4.

[8] Vidgof, M., Bachhofner, S., Mendling, J., Large Language Models for Business Process Management: Opportunities and Challenges. In: Business Process Management Forum. BPM 2023. Lecture Notes in Business Information Processing, vol 490, Springer, Cham. pp. 107–123, 2023, doi: https://doi.org/10.1007/978-3-031-41623-1_7.

[9] Grohs, M., Abb, L., Elsayed, N., Rehse, J.R., Large Language Models Can Accomplish Business Process Management Tasks. In: Business Process Management Workshops. BPM 2023. Lecture Notes in Business Information Processing, vol 492. Springer, Cham, 2023, pp. 453–465. doi: https://doi.org/10.1007/978-3-031-50974-2_34.

[10] Van Der Aalst, W., Data Science in Action. Springer, Second Edition, Berlin Heidelberg, 2016. doi: https://doi.org/10.1007/978-3-662-49851-4.

[11] Gregory, R.P., When is Greenwashing an Easy Fix?, Journal of Sustainable Finance & Investment, 13(2), 2021, pp. 919–942, doi: https://doi.org/10.1080/20430795.2021.1907091.

[12] Minkkinen, M., Niukkanen, A., Mäntymäki, M., What About Investors? ESG Analyses as Tools for Ethics-Based AI Auditing. AI & Society 39(1), 329–343, 2024. doi: https://doi.org/10.1007/s00146-022-01415-0.

[13] Morning Sustainalytics, Official Website, 2025. URL: https://www.sustainalytics.com/.

[14] LSEG, Environmental, Social and Governance Scores from LSEG, 2025. URL: https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf.

[15] Global Reporting Initiative (GRI), Official Website, 2025. URL: https://www.globalreporting.org/.

[16] Sustainability Accounting Standards Board (SASB), Official Website, 2025. URL: https://sasb.org/.

[17] World Economic Forum (WEF), Official Website, 2025. URL: https://www.weforum.org/.

[18] Microsoft, Project GraphRAG – LLM-Derived Knowledge Graphs, 2025. URL: https://www.microsoft.com/en-us/research/project/graphrag/.

[19] Chiş, A., Stoica, O.I., Ghiran, A.M., Buchmann, R.A., A Knowledge Graph Approach to Cyber Threat Mitigation Derived from Data Flow Diagrams. In: 2024 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), pp. 1–6, IEEE, 2024. doi: 10.1109/AQTR61889.2024.10554074

[20] Niculescu, V., Chisalita-Cretu, C., Osman, C.C., Sterca, A., Vector Based Modeling of Business Processes. In: Proceedings of 20th International Conference Evaluation of Novel Approaches to Software Engineering, Science and Technology Publications, Lda, Setúbal, 2024, pp. 735–742. doi: 10.5220/0012739100003687.

[21] Buchmann, R., Eder, J., Fill, H.G., Frank, U., Karagiannis, D., Laurenzi, E., Mylopoulos, J., Plexousakis, D., Santos, M.Y., Large language models: Expectations for semantics-driven systems engineering. Data & Knowledge Engineering 152, 102324, 2024, pp. 1-13. doi: https://doi.org/10.1016/j.datak.2024.102324.

[22] Object Management Group (OMG), Business Process Model and Notation (BPMN) Specification. Version 2.0.2, 2024. URL: https://www.omg.org/spec/BPMN/2.0.2/PDF.

[23] APQC, APQC's Levels of Knowledge Management Maturity, 2025. URL: https://www.apqc.org/resource-library/resource-listing/apqcs-levels-knowledge-management-maturity.

[24] Brinkkemper, S., Saeki, M., Harmsen, F., Assembly Techniques for Method Engineering. In: Advanced Information Systems Engineering. CAiSE 1998. Lecture Notes in Computer Science, vol 1413. Springer, Berlin, Heidelberg, pp. 381–400. doi: https://doi.org/10.1007/BFb0054236.

[25] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems, 24(3), 45–77, 2007. doi: https://doi.org/10.2753/MIS0742-1222240302.

[26] Karagiannis, D., Agile Modeling Method Engineering. In: Proceedings of 19th Panhellenic Conference on Informatics, pp. 5–10, 2015. doi: https://doi.org/10.1145/2801948.2802040.

[27] Ramautar, V., España, S., The OpenESEA Modeling Language and Tool for Ethical, Social, and Environmental Accounting. Complex Systems Informatics & Modeling Quarterly, 34 (2023) 1-29. doi: https://doi.org/10.7250/csimq.2023-34.01.

[28] Karmaker, S.K., Feng, D., TELeR: A General Taxonomy of LLM Prompts for Benchmarking Complex Tasks. In: Findings of the Association for Computational Linguistics: Empirical Methods in Natural Language Processing 2023, pp. 14197–14203. ACL, 2023.

[29] European Institute of Management and Finance (EIMF), The Impact of Artificial Intelligence in Environmental, Social and Governance Reporting, 2025. URL: https://eimf.eu/the-impact-of-artificial-intelligence-in-environmental-social-and-governance-reporting/.

[30] Dydon.ai, Agent AI: A New Era in Sustainable Finance for ESG and EU Taxonomy Reporting, 2025. URL: https://dydon.ai/agent-ai-a-new-era-in-sustainable-finance-for-esg-and-eu-taxonomy-reporting/.

[31] Daněk, J., Pospíšil, J., How Generative AI Autonomous Agents Shape the Current Business Trends and Smart Learning, Trendy v podnikání - Business Trends 2024/2, pp. 35-46, 2024. doi: https://doi.org/10.24132/jbt.2024.14.2.35_46.

[32] Nonaka, I., von Krogh, G., Tacit knowledge and knowledge conversion: controversy and advancement in organisational knowledge creation theory. Organization Science 20 (2009) 635–652. https://doi.org/10.1287/orsc.1080.0412.