Design Science Contributions to the LLM-based Augmentation of the BPM Lifecycle Using Metamodel-based Knowledge Graphs as Domain-specific Context Mediators

Damaris - Naomi Dolha[1,*]

[1]Babeş-Bolyai University, Teodor Mihali 58-60, Cluj-Napoca 400591, Romania

Abstract

Large Language Models (LLMs) have emerged as transformative tools across various domains, including Business Process Management (BPM). Recent research agendas suggest that LLMs offer significant potential to augment the phases of the BPM lifecycle. However, integrating LLMs into BPM tools presents technical choices and challenges, not only with respect to how the process logic is interpreted, but also to potential domain-specific contextualization that may be engineered on top of standard process descriptions.

This PhD work, currently at the end of its first year, investigates the LLM-BPM lifecycle interaction, with a focus on design-time tasks such as process analysis and redesign. This requires the integration of a Business Process Model and Notation (BPMN) tool with public LLM services to enable the experimental probing of content exchanges between the two; moreover, our conceptualization scope will go beyond standard BPMN, involving layers of domain-specificity added to process descriptions - as found in numerous BPMN extensions, and in multi-perspective modeling methods that employ BPMN for the process perspective.

The work will follow the Design Science methodology to implement an experimentation testbed that will enable conversations with LLMs through a BPMN modeling environment - not only about BPMN content, but also pertaining to domain-specific extensions added via metamodeling (DSMLs based on BPMN). Both the BPMN content and domain-specific contexts will be exposed as knowledge graph snippets through a GraphRAG pipeline leveraging diagram-to-RDF converters. In the empirically-focused stages of DSR this tool is intended to help with assessing (a) the performance benefits of the approach against process interpretation by AI visual inspection (computer vision) or by parsing standard XML serializations; (b) the LLM answer sensitivity to metamodeling design decisions (wording and metamodel structure). The ADOxx metamodeling platform will be used for tool and pipeline implementation, and the RAGAs framework will be used to assess the quality of generative content (process queries, modeling recommendations). To probe LLM sensitivity, experiments will vary factors such as prompting strategies, semantic subgraph extraction patterns, metamodeling patterns and terminology.

Keywords

Business Process Management, Context engineering, Domain-specific Modeling, Large Language Models, Graph RAG, Knowledge Graphs

1. Introduction

Organizations increasingly adopt AI approaches to analyze and refine their internal processes, underscoring a need for comprehensive AI-driven frameworks to manage complexity in business process architectures. In this context, [1] defines Business Process Management (BPM) as a holistic discipline that, besides enabling the control, improvement and automation of process models also aligns processes with strategic objectives, thereby enhancing operational efficiency and organizational agility. However, M. Rosemann [2] argues that BPM research must evolve beyond exploitative BPM, which has traditionally focused on optimizing existing processes, towards explorative BPM, which seeks to innovate processes through emerging technologies. In this regard, one wave of transformative advancements is that of Large Language Models (LLMs) opening new avenues - not only for classic Natural Language

ER2025: Companion Proceedings of the 44th International Conference on Conceptual Modeling: Industrial Track, ER Forum, 8th SCME, Doctoral Consortium, Tutorials, Project Exhibitions, Posters and Demos, October 20-23, 2025, Poitiers, France

*Corresponding author.

damaris.dolha@econ.ubbcluj.ro (D. -. N. Dolha)

0009-0001-7695-4383 (D. -. N. Dolha)

Processing (NLP) tasks in relation to BPM, but also for reimagining the BPM lifecycle phases; a recent vision paper [3] applied this lens to revisit the lifecycle and proposed a research agenda that inspired this thesis work.

Besides empirically probing the ability of public LLMs to interpret process descriptions, the integration of LLMs into BPM tools also raises artifact-oriented challenges that must be tackled through a Design Science Research (DSR) approach - i.e. investigating design prescriptions for schemas, formats, protocols and knowledge structures that should be used for streamlining content between an LLM and a BPMN tool, for domain-specific context engineering and for assessing the sensitivity of LLM response quality to various content strategies and context patterns.

Considering this requirement for engineering-empirical duality, this PhD work aims to develop a twofold Design Science contribution:

(a) Firstly, a testbed (modeling tool) to facilitate experimentation on the above aspects, through a BPMN-based DSML (domain-specific modeling language) that captures not only standard BPMN descriptions but also domain-specific extensions beyond the standard, editable by metamodeling means; interoperation between this modeling tool and LLMs will be based on RDF graphs to enable a model-driven Graph RAG (Graph Retrieval-Augmented Generation) pipeline that will support, via RDF fragments, process-awareness and domain-awareness for improving LLM answers quality and relevance in domain-specific terms;

(b) Secondly, to employ this testbed for experimentation with bidirectional conversations (i.e. process queries and modeling suggestions). The planned experiments aim to evaluate the performance benefits of this model-driven Graph RAG pipeline against a baseline of common process inspection approaches - visual diagram inspection (via computer vision associated with some LLM services) or ingestion of standard BPMN XML serializations annotated with unstructured domain-specific context documents. To formulate prescriptive knowledge as expected in DSR, the experimentation strategy aims to assess the sensitivity of LLM responses to various design decisions and interoperability strategies involved in the Graph RAG pipeline: the metamodel terminology, metamodel design patterns, prompting strategies, "relevant context" subgraph extraction strategies.

The intrinsic gap between the unstructured and probabilistic outputs of LLMs and the rigorously structured demands of business process modeling standards necessitates reconciliation approaches and a good understanding of their limits. Although LLMs can be instructed using different strategies to generate and analyze conceptual models, their outputs often require some post-processing to ensure adherence to formal modeling conventions [4]. While prompt engineering alone can leverage pre-trained LLMs without any fine-tuning, just by wording choices and prompting protocols - it has limitations in fully communicating formal syntax and semantics of established BPM standards [5]. To address this, recent research agendas [6] highlighted the need for LLM integration with semantic structuring mechanisms - KGs, upper ontologies or reasoning rulebases - to ensure that process interpretability is not compromised by the probabilistic nature of LLMs. Therefore, the interplay between LLMs and the BPM lifecycle calls for new integration patterns and hybridization architectures. As argued by [7], a hybrid approach seems more promising than solely training chatbots with specialized process repositories, preferably leveraging existing modeling knowledge.

In this light, this doctoral research - currently at the end of its first year - aims to investigate **to what extent can LLM-powered augmentations support specific phases of the BPM lifecycle, by leveraging knowledge graphs as conceptual mediators?** To delimit a realistic scope for the work, we focus on the phases of: *Process analysis* - via process queries and simulation; and *Process (re)design* - via domain-specific modeling recommendations. This scope is further refined into research questions in the next section.

2. Research Questions

Considering the scoping established above, several research questions are formulated below based on the DSR taxonomy of RQs proposed by [8]:

RQ1 (artifact-focused). How can we implement LLM-powered process analysis in a domain-specific BPMN tool, so that *the LLM can answer both in terms of the standard BPMN conceptualization and the non-standard domain specificity added to BPMN*? This will go beyond the current investigation of LLMs ability to enable conversation with BPMN-compliant content, towards exploring whether that capability extends to open-ended domain-specific contextualization of process descriptions (i.e., DSMLs based on BPMN) in terms of various semantic patterns - new taxonomies, new attributes, new relationships to new concepts. Many process-centric DSMLs are BPMN extensions realized by well established means - by enriching the taxonomies of tasks, flows, document objects etc. (see some examples in [9, 10, 11, 12]), by adding entirely new model types that are meaningfully connected to legacy BPMN [13]. It cannot be expected that such metamodel extensions are available in the training phase of commercial or public LLMs therefore domain-specificity must be added by hybridization mechanisms - such as RAG or ontology-guided prompting [14]. We aim to distance this work from the existing approach proposed in the BPMN-LLM framework of [15] and instead leverage as a technical ingredient semantic graphs (obtained from diagram patterns [16]) and the *Graph RAG* integration patterns[1] generally reported as superior to traditional chunking-based RAG [17], since they employ semantic strategies for context engineering and extraction.

This technical decision is founded on the possibility of capturing the domain-specificity of a DSML through KGs obtained from the RDF enrichment of standard BPMN process diagrams. Even if not yet supported by a standard, such BPMN-to-RDF convertors have been available in several experimental or didactic modeling tools for some years now [18, 19]. Our demonstrator will employ the ADOxx metamodeling platform[2] to extend the Bee-Up implementation of BPMN. Therefore, the engineering effort will leverage semantic graphs as mediators between a BPMN-based DSML and LLM services, thus providing a technology-specific operationalization of the "mediator role" of conceptual models advocated by recent agendas in conceptual modeling research [20] and semantics-driven systems engineering [6]. Besides acting as a demonstrator for the proposed DSR treatment, the outcome of this RQ will further provide a testbed for the experimentation oriented RQs:

RQ2 (empirically-focused). How does this implementation based on Graph RAG perform in comparison to more straightforward alternatives - i.e., an AI agent performing visual diagram inspection, or ingesting standard XML process serializations? In preliminary results, the first year of this PhD project already published some results in [21, 22] focusing on comparisons of LLM answer quality between BPMN content provided as standard XML and similar content provided as tool-specific RDF graphs, using the OpenAI integration of Ontotext GraphDB[3] as interoperability channel. Results suggest a benefit of open-endedness and relationship traceability in RDF graphs compared to the closed-world assumption of standard XML schemas.

RQ3 (empirically-focused). How do metamodeling design decisions - from terminological decisions (i.e., wording of metamodel constructs) to knowledge modeling decisions - impact the LLM ability to interpret and reason on business processes in the domain-specific context of choice?

Regarding the **domain-specificity** involved in all the above RQs - i.e., the application domain for the DSR demonstrator - there are numerous choices of BPMN extensions available in the literature, see [9, 10]. Since we require open source implementations that can be both edited (for RQ3) and extended with interoperability components (for RQ2), we are leveraging the open resources made available through the OMILAB Community of Practice [23], where a large catalog of domain-specific modeling tools have been implemented over the years and are editable on the ADOxx metamodeling environment, including numerous domain-specific flavors of BPMN (as needed for our RQ1). Out of the numerous options, our current focus is on a **BPMN extension for ESG** (Environmental-Social-Governance)

[1] https://www.ontotext.com/knowledgehub/fundamentals/what-is-graph-rag/
[2] https://adoxx.org/
[3] https://graphdb.ontotext.com/

policies [11] - a wider tooling project[4] spanning several PhD projects, where this PhD also participates in the domain analysis and tool development efforts. Our results are however expected to be generalizable and will also be tested with alternative specificities, to cover a diversity of metamodel and diagram patterns that have been catalogued in past works in both functional and formal terms [24, 25].

3. Literature Survey on LLMs for BPM

A recent manifesto [26] extends traditional Business Process Management Systems (BPMS) by integrating AI into every phase of the lifecycle, resulting in AI-Augmented Business Process Management Systems (ABPMSs). Unlike conventional BPM - where tasks are executed by humans or software following predetermined logic - ABPMSs are supposed to manifest some level of autonomy in deciding task execution. To mitigate potential trust concerns for users due to AI's opaque decision-making processes, the authors envision a lifecycle that distinguishes between basic (frame, enact, perceive, reason) and advanced phases (explain, adapt, improve), aimed at making AI decisions more transparent and understandable. This is in line with similar preoccupations in the systems engineering field, e.g., the hybridization in semantics-driven systems engineering [6]. The earlier agenda of [3] pinpoints how LLMs impact BPM, from process identification and discovery to monitoring, and how they are expected to be integrated into commercial products. The proposal of [7] narrows the focus on three phases of the BPM lifecycle - namely, process discovery, analysis and redesign - examining how chatbots can support conversational process modeling. Their results show that chatbots perform well in assisting with task extraction, paraphrasing and refining process descriptions, although they still face challenges in fully "understanding" process logic and precise meaning from process descriptions. Such developments have led to a growing body of work exploring the use of LLMs in distinct areas of BPM practice, notably process generation, process querying, process discovery. The following section indicates some representative contributions and challenges in each of these segments, followed by a shift to our focus on multi-modal knowledge streamlining - visual, knowledge graphs and prompt engineering.

3.1. Prominent BPM Use Cases Employing LLMs

Process querying. An important use case of LLMs within the BPM paradigm is generating natural language explanations in response to user queries about their business processes. This tends to replace the tradition of process querying methods [27], in similar sense to how various querying standards (Cypher, SPARQL, SQL) are being considered as mediators for natural language inquiries. In this regard, [15] proposes and evaluates a RAG and LLM fine-tuning architecture, devoting a significant portion of their study to the effect of different chunking strategies on the accuracy of LLM-generated outputs. Recent work also explores how generative chatbots can support process querying by interactively guiding users through BPMN models. For instance, [28] reports that LLM-powered chatbots can answer questions about available tasks and process decisions, even though they struggle with complex control flows and gateways. When it comes to trustworthiness and clarity of AI-generated explanations, [29] demonstrates that adding process and causal knowledge to LLM prompts improve the results of such metrics. Explainability is also present from the very first prompt in [30], where ChatGPT[5] (using GPT-4o) is asked to describe uploaded BPMN diagrams - before any other steps are taken - noting that ChatGPT has the ability to visually inspect process models and recognize the BPMN symbols. In our work, we stay away from visual communication of process designs - first of all, because not all business process details manifest visually; secondly, because BPMSs rely on model interchange for deterministic communication of design decisions to inform use cases where stochastic extrapolation is not acceptable (e.g., automation); finally, because arbitrary domain-specificity added to BPMN would not be recognized visually, while having the same domain-specificity in semantic graphs makes it interpretable.

[4] presented in the Research Project Exhibition at CAiSE 2025: https://github.com/claudenirmf/caise-rpe-2025/blob/main/caise-rpe-01.pdf

[5] https://openai.com/index/chatgpt/

Process generation. From extracting process elements and relationships directly from textual narratives [31], recent research also demonstrates that even early LLMs versions (such as OpenAI's GPT-3.5/4) have the potential to generate and even refine, in certain cases, business process diagrams from textual descriptions [4, 32, 33, 34, 35] and voice inputs [36]. Nonetheless, some outputs require external/human refinement to meet semantic correctness and adherence to modeling conventions, a challenge also noticed in the framework introduced by [37]. Here, the authors further note the need for more interactive feedback mechanisms to improve the generated BPMN models. [38] reinforces the idea that LLMs alone are insufficient for process modeling in a context-agnostic setting. This study highlights that an effective use of these technologies requires a governance structure, where human process analysts and AI systems iteratively refine each other's work. An LLM-powered chatbot is integrated with an existing modeling environment, pointing to the need of setting up testbeds with well-identified points for tweaking factors that potentially impact experimentation outcomes - a mission toward which this thesis also aligns through its DSR approach. This thesis does not pursue the goal of generating standards-compliant business process models, but rather to assist modelers through recommendations of domain-specific elements to be attached to process models, considering the adopted domain-specificity represented by a BPMN-centric DSML.

Process discovery. At the base of data-driven BPM, [39] highlights that process mining is evolving toward AI-augmented process execution. According to [40], Large Process Models (LPM) - essentially neuro-symbolic, AI-augmented systems - reconceptualize process mining as an interactive, human-in-the-loop dialogue mediated by natural language, rather than a static, tool-driven extraction activity. Preliminary experiments with GPT-4 (where process models are created from operational data) show that LLMs can interpret process data, answer queries, detect anomalies and suggest improvements through natural language [41]. Within this phase, major concerns are detecting anomalies and fairness assessments [42], with results supporting the idea that commercial LLMs are generally capable to support analytics tasks involved in process discovery. This thesis does not aim currently to pursue a process mining research stream, although we are open to the possibility that the modeling assistance recommendation use case may branch out towards assessing the analytics capabilities of LLMs.

3.2. Prompt Engineering and Knowledge Graphs in BPM

In order to get the best of LLMs within BPM, studies such as [43, 44, 45] confirm the critical role of prompt engineering, particularly in complex tasks requiring structured responses. Prompt engineering was investigated for BPM applications [45] to improve semantic completeness in process models - and in the context of KGs [46] to enhance knowledge extraction and reasoning. Several challenges remain open to investigation, especially with the fast evolution of LLM capabilities: (a) prompt sensitivity, where minor variations in wording significantly alter outputs [44, 46] - an approach where this thesis will compare the sensitivity to terminological choices and schemas governing graphical process models; (b) domain adaptation difficulties, as general-purpose prompting struggles to capture nuance and requires specialized strategies - an issue investigated in this work by a metamodeling approach to injecting domain-specificity to process descriptions; and (c) prompting strategy complexity [45].

These challenges underscore that instruction-based prompts alone are insufficient in ensuring reliable LLM-generated outputs. Knowledge injection techniques have been proposed as a means to improve assertions while decreasing hallucinated outputs [47]. Moreover, in [32], the Process Knowledge Graph serves as a centralized, dynamic repository of process-related knowledge, for context-aware workflow generation, ensuring that AI-driven process recommendations align with evolving organizational requirements. The paper [48] further emphasizes the need for structured semantic grounding through KGs. However, to maximize the effectiveness of KGs, it is essential to filter them by subgraph extraction mechanisms to build relevant context patterns - the presence of noisy triples can mislead LLMs even when effective prompting techniques are used [49]. A popular Graph RAG pattern employs metadata filtering based on user queries, so that only pertinent information is retrieved [50]. The integration

of AI-driven functionalities in prominent BPM commercial tools like ADONIS[6] and SAP Signavio[7] exemplifies industry-grade approaches that rely on fine-tuning, whereas we aim to investigate the inner workings of LLM-BPM tool interactions with KGs as mediators, potentially enabling an open-endedness to domain-specificity added via metamodeling.

4. Research Methodology

The DSR process [51, 52] is particularly suited for artifact-oriented research. A specific tuning of it will manifest in the *Design & Development* phase which will be delegated to a traditional modeling tool engineering methodology - AMME (Agile Modeling Method Engineering) [53]. This is due to the intended use of the ADOxx metamodeling platform to add domain-specificity to the BPMN implementation available in the Bee-Up tool[8] and of the ADOxx-to-RDF diagram converters available from previous projects [16]. Figure 1 shows the architectural vision of the model-driven Graph RAG pipeline, with (a) the Bee-Up extension as a domain-specific enrichment of BPMN (the ESG-flavored implementation in [11] being a primary candidate in the initial DSR iteration) and (b) Python orchestrating the Graph RAG pipeline elements, with Ontotext GraphDB and OpenAI as KG store and as LLM service, respectively. While a fast evolution of LLM services is expected, the domain-specificity of choice will always be a non-standard conceptualization, while BPMN is already fairly known to public LLMs.

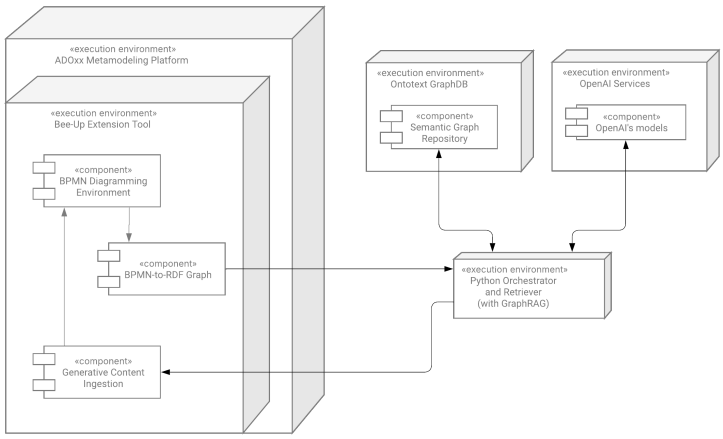

Figure 1: Architectural vision of the experimentation pipeline.

In the *Evaluation* phase, this work will contribute a protocol for evaluating the LLM response/content quality relative to ground truths and in relation to several varying design factors - metamodel patterns and terminology, prompting complexities and subgraph extraction strategies. Correctness criteria will be based on the RAGAs [54] framework, which defines computable metrics such as *Response Relevancy* (how well the answer aligns with the user's question intent while penalizing irrelevant information), *Factual Correctness* (factual overlap against a ground truth) and others. Performance and cost of the pipeline relative to OpenAI or similar services will also be considered as part of the evaluation. There is no plan to evaluate the actual DSML involved in experimentation, since the thesis is not a language/notation development project. We are interested in building a domain-specific context engineering flow between visual process modeling and LLM agents/services, and domains will be switched for comparative experimentation of semantic patterns, to formulate a notion of "LLM readiness" for BPMN-based languages.

[6] https://www.boc-group.com/en/blog/bpm/adonis-and-ai/
[7] https://www.signavio.com/process-ai/
[8] https://bee-up.omilab.org/activities/bee-up/

5. Preliminary Efforts

The first steps developed in the first year of the PhD program [21, 22] looked at how OpenAI's 2024 models (GPT 3-4) interpreted BPMN models ingested as RDF graphs exported from Bee-Up (via Ontotext GraphDB) versus standard XML diagrams exported from the Signavio toolkit. Experimental results suggest that semantic graphs are more suitable to be interpreted and navigated along chains of relationships even when using a non-standard terminology, possibly due to the intricate network of XML cross-references involved in answering questions that must combine navigation of connectors, containments, data annotations or inter-model hyperlinks (e.g., subprocess links). RDF representations as Turtle "sentences" more effectively preserve all properties, whereas with XML, the generative content is more prone to hallucinate, even extrapolating business narratives from labels present in diagrams. Experiments were however limited in design conditions and imprecise in concern separation, as well as in evaluating the LLM response quality, while LLMs are quickly evolving and improving on all evaluated aspects. During such experimentation efforts we arrived at the metrics offered by the RAGAs framework (and Python library for computing them) - to be further involved in measuring the quality of generative outputs. A more systematic assessment protocol will evolve from these early steps, also incorporating the variation factors we are interested in - i.e., sensitivity of LLM responses to metamodeling design decisions (wording/terminology as well as metamodel patterns) and communication strategies (prompting structure, the subgraph patterns around a graph node of interest as a replacement for "chunks" in traditional RAG). Prompting strategies are being designed along the TELeR prompt taxonomy [43] to probe the capabilities of reasoning and traceability, first limited to the BPMN standard, then extended with non-standard domain-specific taxonomies, looking towards managing BPMN repositories as knowledge graphs, a recent proposition of [18] and [16].

In terms of content subjected to the GraphRAG pipeline, current work focused only on basic BPMN for which we defined a collection of minimalist workflow patterns involving a diversity of relationships present in BPMN, considering most used BPMN usage trends from [55]. In parallel, the ESG extension to BPMN reported in [11] started development, as a distinct DSML engineering project to be adopted as an application case in the future work (as also suggested by the report in the Research Project Exhibition of CAiSE 2025: https://github.com/claudenirmf/caise-rpe-2025/blob/main/caise-rpe-01.pdf).

6. Future Work

This work follows a DSR process to build a knowledge pipeline between BPMN environments and LLM services, with a focus on how domain-specific process contextualizations can contribute to that interaction. Instead of relying on AI capabilities for visual inspection or for interpretation of XML standard serializations, we are moving to semantic graphs as a mediator between metamodel-enriched BPMN and LLMs interpreting those enrichments. For this, the development of a Graph RAG pipeline to ensure streamlined interoperability between a metamodeling environment and an LLM is a priority of the next phase of the project. This will leave the second half of the PhD program for experimentation and benchmarking of LLMs' generative content quality in relation to various design decisions of BPMN extensions, and how they are exposed to LLMs in use cases based on querying or modeling assistance.

Acknowledgements

This PhD work is supervised by Prof. Dr. Robert Andrei Buchmann, at University Babeş-Bolyai, Romania, in the Information Systems domain.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] M. Dumas, M. L. Rosa, J. Mendling, H. A. Reijers, Fundamentals of Business Process Management, 2nd ed., Springer, 2018. doi:10.1007/978-3-662-56509-4.

[2] T. Kohlborn, O. Mueller, J. Poeppelbuss, M. Roeglinger, Interview with Michael Rosemann on ambidextrous business process management, Bus. Proc. Manage. J. 20 (2014) 634–638. doi:10.1108/BPMJ-02-2014-0012.

[3] M. Vidgof, S. Bachhofner, J. Mendling, Large Language Models for Business Process Management: Opportunities and Challenges, in: Proceedings of BPM 2023 Forum, volume 490 of *LNBIP*, Springer, 2023, pp. 107–123. doi:10.1007/978-3-031-41623-1_7.

[4] H.-G. Fill, P. Fettke, J. Köpke, Conceptual Modeling and Large Language Models: Impressions From First Experiments With ChatGPT, EMISAJ 18 (2023). doi:10.18417/emisa.18.3.

[5] K. Busch, A. Rochlitzer, D. Sola, H. Leopold, Just Tell Me: Prompt Engineering in Business Process Management, in: Proceedings of BPMDS EMMSAD 2023, volume 479 of *LNBIP*, Springer, 2023, pp. 3–11. doi:10.1007/978-3-031-34241-7_1.

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

[7] N. Klievtsova, J. V. Benzin, T. Kampik, J. Mangler, S. Rinderle-Ma, Conversational Process Modelling: State of the Art, Applications, and Implications in Practice, in: Proceedings of BPM 2023 Forum, volume 490 of *LNBIP*, Springer, 2023, pp. 319–336. doi:10.1007/978-3-031-41623-1_19.

[8] N. H. Thuan, A. Drechsler, P. Antunes, Construction of Design Science Research Questions, Communications of the AIS 44 (2019). doi:10.17705/1CAIS.04420.

[9] F. Corradini, F. Fornari, S. Pettinari, B. Re, L. Rossi, F. Tiezzi, A BPMN-Based Approach for IoT Systems Engineering, in: Fluidware, Internet of Things, Springer, 2024, pp. 85–105. doi:10.1007/978-3-031-62146-8_5.

[10] A. Chiș, A.-M. Ghiran, BPMN Extension for Multi-Protocol Data Orchestration, in: Domain-Specific Conceptual Modeling, Springer, 2022, pp. 639–656. doi:10.1007/978-3-030-93547-4_28.

[11] C.-C. Osman, A.-M. Ghiran, R. A. Buchmann, Towards a Knowledge Management Capability for ESG Accounting with the Help of Enterprise Modeling and Knowledge Graphs, in: Proceedings of PoEM 2024 co-located events, CEUR-WS, 2024. URL: https://ceur-ws.org/Vol-3855/forum12.pdf.

[12] Ș. Uifălean and R. A. Buchmann, Low-Code Browser Front-End Automation Using RDF Graphs and a Domain-Specific Language for UX Representation, in: Proceedings of RCIS 2025, volume 547 of *LNBIP*, Springer, 2025, pp. 140–155. doi:10.1007/978-3-031-92474-3_9.

[13] A. Chiș, A Modeling Method for Work Systems Knowledge Capture and Traceability, in: Proceedings of CAiSE 2025, volume 557 of *LNBIP*, Springer, 2025, pp. 239–246. doi:10.1007/978-3-031-94590-8_29.

[14] V. I. Iga, G. C. Silaghi, Ontology-Based Dialogue System for Domain-Specific Knowledge Acquisition, in: Proceedings of ISD 2023, AIS, 2023. doi:10.62036/ISD.2023.46.

[15] M. L. Bernardi, A. Casciani, M. Cimitile, A. Marrella, Conversing with Business Process-Aware Large Language Models: The BPLLM Framework, J. Intell. Inf. Syst. 62 (2024) 1607–1629. doi:10.1007/s10844-024-00898-1.

[16] M. Cinpoeru, A.-M. Ghiran, A. Harkai, R. A. Buchmann, D. Karagiannis, Model-Driven Context Configuration in Business Process Management Systems: An Approach Based on Knowledge Graphs, in: Proceedings of BIR 2019, volume 365 of *LNBIP*, Springer, 2019, pp. 189–203. doi:10.1007/978-3-030-31143-8_14.

[17] T. C. Nemțoc, A.-M. Ghiran, Natural Language Querying of Invoice Data Using RAG and GraphRAG: Leveraging LLMs for Financial Document Insights, in: Proceedings of CAiSE 2025 Workshops, volume 556 of *LNBIP*, Springer, 2025, pp. 69–80. doi:10.1007/978-3-031-94931-9_6.

[18] S. Bachhofner, E. Kiesling, K. Revoredo, P. Waibel, A. Polleres, Automated process knowledge graph construction from BPMN models, in: Proceedings of DEXA 2022, volume 13426 of *LNBIP*, Springer, 2022, pp. 32–47. doi:10.1007/978-3-031-12423-5_3.

[19] OMiLAB NPO, The Bee-Up Modeling Tool, n.d. URL: https://bee-up.omilab.org/activities/bee-up/.
[20] J. Recker, R. Lukyanenko, M. Jabbari, B. M. Samuel, A. Castellanos, From Representation to Mediation: A New Agenda for Conceptual Modeling Research in a Digital World, MIS Quarterly 45 (2021) 269–300. doi:10.25300/MISQ/2021/16027.
[21] D. N. Dolha, R. A. Buchmann, Experiments with natural language queries on RDF vs. XML-serialized BPMN diagrams, in: Proceedings of KES 2024, volume 246, 2024, pp. 3246–3255. doi:10.1016/j.procs.2024.09.315.
[22] D. N. Dolha, R. A. Buchmann, Generative AI for BPMN Process Analysis: Experiments with Multi-modal Process Representations, in: Proceedings of BIR 2024, volume 529 of *LNBIP*, Springer, 2024, pp. 19–35. doi:10.1007/978-3-031-71333-0_2.
[23] The OMiLAB Community, Development of Conceptual Models and Realization of Modelling Tools Within the ADOxx Meta-Modelling Environment: A Living Paper, in: Domain-Specific Conceptual Modeling, Springer, 2022, pp. 23–40. doi:10.1007/978-3-030-93547-4_2.
[24] R. A. Buchmann, D. Karagiannis, Pattern-based Transformation of Diagrammatic Conceptual Models for Semantic Enrichment in the Web of Data, in: Proceedings of KES 2015, volume 60, 2015, pp. 150–159. doi:10.1016/j.procs.2015.08.114.
[25] H.-G. Fill, T. Redmond, D. Karagiannis, Formalizing Meta Models with FDMM: The ADOxx Case, in: Proceedings of ICEIS 1012, volume 141 of *LNBIP*, Springer, 2013, pp. 429–451. doi:10.1007/978-3-642-40654-6_26.
[26] M. Dumas, et al., AI-augmented Business Process Management Systems: A Research Manifesto, ACM Trans. Manage. Inf. Syst. 14 (2023). doi:10.1145/3576047.
[27] A. Polyvyanyy, Process Querying: Methods, Techniques, and Applications, in: Process Querying Methods, Springer, 2022, pp. 511–524. doi:10.1007/978-3-030-92875-9_18.
[28] L. F. Lins, N. Nascimento, P. Alencar, T. Oliveira, D. Cowan, Comparing Generative Chatbots Based on Process Requirements: A Case Study, in: Proceedings of IEEE Big Data 2023, IEEE, 2023, pp. 4664–4673. doi:10.1109/BigData59044.2023.10386251.
[29] D. Fahland, F. Fournier, L. Limonad, I. Skarbovsky, A. J. E. Swevels, How well can a large language model explain business processes as perceived by users?, Data & Knowledge Engineering 157 (2025) 102416. doi:10.1016/j.datak.2025.102416.
[30] V. Niculescu, M.-C. Chisăliță-Crețu, C.-C. Osman, A. Sterca, Model-Driven Development Using LLMs: The Case of ChatGPT, in: Proceedings of ENASE 2025, SciTePress, 2025, pp. 328–339. doi:10.5220/0013484400003928.
[31] P. Bellan, M. Dragoni, C. Ghidini, Extracting Business Process Entities and Relations from Text Using Pre-trained Language Models and In-Context Learning, in: Proceedings of EDOC 2022, volume 13585 of *LNCS*, Springer, 2022, pp. 182–199. doi:10.1007/978-3-031-17604-3_11.
[32] A. Beheshti, J. Yang, Q. Z. Sheng, B. Benatallah, F. Casati, S. Dustdar, H. R. M. Nezhad, X. Zhang, S. Xue, ProcessGPT: Transforming Business Process Management with Generative Artificial Intelligence, in: Proceedings of ICWS 2023, IEEE CS, 2023, pp. 731–739. doi:10.1109/ICWS60048.2023.00099.
[33] H. Kourani, A. Berti, D. Schuster, W. M. P. van der Aalst, ProMoAI: Process Modeling with Generative AI, in: Proceedings of IJCAI 2024, ACM Press, 2024, pp. 8708–8712. doi:10.24963/ijcai.2024/1014.
[34] J. Silva, Q. Ma, J. Cabot, P. Kelsen, H. A. Proper, Application of the Tree-of-Thoughts Framework to LLM-Enabled Domain Modeling, in: Proceedings of ER 2024, volume 15238 of *LNCS*, Springer, 2025, pp. 94–111. doi:10.1007/978-3-031-75872-0_6.
[35] M. Grohs, L. Abb, N. Elsayed, J. R. Rehse, Large Language Models Can Accomplish Business Process Management Tasks, in: Proceedings of BPM 2023 Workshops, volume 492 of *LNBIP*, Springer, 2024, pp. 453–465. doi:10.1007/978-3-031-50974-2_34.
[36] J. Köpke, A. Safan, Introducing the BPMN-Chatbot for Efficient LLM-Based Process Modeling, in: Proceedings of BPM 2024 co-located events, CEUR-WS, 2024. URL: https://ceur-ws.org/Vol-3758/paper-15.pdf.
[37] H. Kourani, A. Berti, D. Schuster, W. M. P. van der Aalst, Process Modeling with Large Language

Models, in: Proceedings of BPMDS EMMSAD 2024, Springer, 2024, pp. 229–244. doi:10.1007/978-3-031-61007-3_18.
[38] C. Ziche, G. Apruzzese, LLM4PM: A Case Study on Using Large Language Models for Process Modeling in Enterprise Organizations, in: Proceedings of BPM 2024 Forum, volume 527 of *LNBIP*, Springer, 2024, pp. 472–483. doi:10.1007/978-3-031-70445-1_35.
[39] D. Chapela-Campa, M. Dumas, From process mining to augmented process execution, Software and Systems Modeling 22 (2023) 1977–1986. doi:10.1007/s10270-023-01132-2.
[40] T. Kampik, et al., Large Process Models: A Vision for Business Process Management in the Age of Generative AI, Künstliche Intelligenz 39 (2025) 81–95. doi:10.1007/s13218-024-00863-8.
[41] A. Berti, D. Schuster, W. M. P. van der Aalst, Abstractions, Scenarios, and Prompt Definitions for Process Mining with LLMs: A Case Study, in: Proceedings of BPM 2023 Workshops, volume 492 of *LNBIP*, Springer, 2024, pp. 427–439. doi:10.1007/978-3-031-50974-2_32.
[42] A. Berti, H. Kourani, W. M. P. van der Aalst, PM-LLM-Benchmark: Evaluating Large Language Models on Process Mining Tasks, in: Proceedings of ICPM 2024 Workshops, volume 533 of *LNBIP*, Springer, 2025, pp. 610–623. doi:10.1007/978-3-031-82225-4_45.
[43] S. K. Karmaker, D. Feng, TELeR: A General Taxonomy of LLM Prompts for Benchmarking Complex Tasks, in: Proceedings of EMNLP 2023, ACL, 2023, pp. 14197–14203. URL: https://aclanthology.org/2023.findings-emnlp.946.pdf.
[44] S. Schulhoff, et al., The Prompt Report: A Systematic Survey of Prompt Engineering Techniques (2024). arXiv:2406.06608.
[45] S. Ayad, F. Alsayoud, Prompt engineering techniques for semantic enhancement in business process models, Bus. Proc. Manage. J. 30 (2024) 2611–2641. doi:10.1108/BPMJ-02-2024-0108.
[46] V. I. R. Iga, G. C. Silaghi, LLMs for Knowledge-Graphs Enhanced Task-Oriented Dialogue Systems: Challenges and Opportunities, in: Proceedings of CAiSE 2024 Workshops, Springer, 2024, pp. 168–179. doi:10.1007/978-3-031-61003-5_15.
[47] A. Martino, M. Iannelli, C. Truong, Knowledge Injection to Counter Large Language Model (LLM) Hallucination, in: Proceedings of ESWC 2023 Satellite Events, Springer, 2023, pp. 182–185. doi:10.1007/978-3-031-43458-7_34.
[48] B. Reitemeyer, H.-G. Fill, Applying Large Language Models in Knowledge Graph-based Enterprise Modeling: Challenges and Opportunities (2025). arXiv:2501.03566.
[49] Q. Zhang, J. Dong, H. Chen, D. Zha, Z. Yu, X. Huang, KnowGPT: Knowledge Graph based Prompting for Large Language Models, in: Proceedings of NIPS 2024, volume 37, Curran Associates Inc., 2025, pp. 6052–6080. URL: https://neurips.cc/virtual/2024/poster/95299.
[50] GraphRAG, GraphRAG Metadata Filtering, 2024. URL: https://graphrag.com/reference/graphrag/metadata-filtering/.
[51] R. J. Wieringa, Design Science Methodology for Information Systems and Software Engineering, Springer, 2014. doi:10.1007/978-3-662-43839-8.
[52] K. Peffers, T. Tuunanen, M. A. Rothenberger, S. Chatterjee, A Design Science Research Methodology for Information Systems Research, J. Manage. Inf. Syst 24 (2007) 45–77. doi:10.2753/MIS0742-1222240302.
[53] D. Karagiannis, Agile modeling method engineering, in: Proceedings of PCI 15, ACM Press, 2015, pp. 5–10. doi:10.1145/2801948.2802040.
[54] S. Es, J. James, L. Espinosa Anke, S. Schockaert, RAGAs: Automated evaluation of retrieval augmented generation, in: Proceedings of EACL 2024, ACL, 2024, pp. 150–158. doi:10.18653/v1/2024.eacl-demo.16.
[55] I. Compagnucci, F. Corradini, F. Fornari, B. Re, A Study on the Usage of the BPMN Notation for Designing Process Collaboration, Choreography, and Conversation Models, Bus. Inf. Syst. Eng. 66 (2024) 43–66. doi:10.1007/s12599-023-00818-7.