Technical, Organizational and Ethical Aspects of Integrating Enterprise Large Language Models into Existing Information Systems Infrastructure*

Lara Kuhlmann[1,*], Alexander van der Staay[2], Maximilian Nebel[2], Annika Zemke[1], Christian Janiesch[2] and Michael Henke[1]

[1]Chair of Enterprise Logistics, TU Dortmund University, Leonhard-Eueler-Straße 5, 44227 Dortmund, Germany
[2]Chair of Enterprise Computing, TU Dortmund University, Otto-Hahn-Str. 12, 44227 Dortmund, Germany

Abstract

Generative AI (GenAI) is becoming increasingly important and many organizations are implementing privately deployed enterprise LLMs. However, the requirements for integrating such systems into existing information systems (IS) infrastructures remain insufficiently understood. This study addresses this gap through an in-depth case study. We analyze the integration of enterprise LLMs from a holistic socio-technical perspective, focusing on technical, organizational, and ethical aspects. The findings show that successful integration cannot be reduced to a purely technical challenge but depends on the joint alignment of technological, organizational, and ethical requirements. In particular, our results highlight the importance of explainability, transparency, and safety properties as key factors that support the responsible and effective integration of enterprise LLMs. Based on these insights, we derive operationalization measures that provide actionable guidance for deploying enterprise LLMs within existing organizational infrastructures.

Keywords

Enterprise Large Language Models, Integration, TOE Framework

1. Introduction

The adoption of generative artificial intelligence (GenAI) is rapidly expanding in organizational contexts. The rapid rise of cloud-based AI as a Service (AIaaS) offerings, and cloud-deployed large language models (LLMs) in particular, has lowered access barriers and enabled rapid adoption across many use cases [1, 2, 3]. At the same time, reliance on external cloud services raises concerns regarding the protection of intellectual property and data privacy [4, 5]. Consequently, organizations face the challenge of leveraging the innovative potential of GenAI while maintaining oversight and complying with regulatory requirements [6]. In response to these challenges, many organizations have begun implementing privately deployed *enterprise LLMs*. Prior research shows that such enterprise AI solutions can support everyday work activities which require enterprise-internal knowledge, for example by handling human resources inquiries, or summarizing internal documents [7]. These benefits are unlocked via fine-tuning of LLMs for domain- or task-specific applications, using enterprises' own data sources [8].

However, integrating an enterprise LLM presents novel technological, organizational, and ethical challenges across the entire system lifecycle. While prior research has examined the capabilities and impacts of GenAI [1] as well as the integration of AIaaS into externalized service offerings [3, 9], the intricacies of enterprise LLM integration remain insufficiently explored. To address this gap, this study investigates the integration of enterprise LLMs from a holistic socio-technical perspective by addressing the following research questions:

RQ1: *What are the requirements for integrating enterprise LLMs into existing IS infrastructures?*

1st Workshop on Explainability, Transparency, and Safety (ExTraSafe)
*Corresponding author.
lara.kuhlmann@tu-dortmund.de (L. Kuhlmann)
0009-0006-2934-0203 (L. Kuhlmann); 0009-0009-6104-1512 (A. v. d. Staay); 0009-0002-9525-6561 (M. Nebel); 0000-0002-8050-123X (C. Janiesch); 0000-0002-5848-9940 (M. Henke)

RQ2: *How can these requirements be operationalized across technological, organizational, and ethical dimensions?*

To address these questions, we conducted an in-depth single-case study using the lens of the technology, organization and ethics (TOE) framework [10]. The study contributes to research on enterprise AI by providing empirical insights into the architectural and governance-related requirements for integrating enterprise LLMs. We also highlights how explainability, transparency, and safety properties must be addressed at a broader socio-technical system.

2. Background

The integration of enterprise LLMs poses a key distinction from the use of publicly accessible systems such as ChatGPT, Claude and Mistral. by leveraging LLMs for secure, customized, and domain-specific applications. Rather than relying on public cloud-based LLM services, many organizations deploy proprietary or open-source models within on-premises or private cloud environments. This approach enables organizations to address specific business requirements while maintaining control over data protection, regulatory compliance, and system customization [11]. Overall, such deployments should aim for safe implementation. In this context, safety refers to preventing unintended harms resulting from enterprise LLM use.

Several key motivations drive the adoption of enterprise LLM solutions, which can be understood along with the TOE dimensions. First, data security and privacy are critical technological and ethical concerns, as enterprises often process sensitive or proprietary information that cannot be exposed to external providers. A key promise of enterprise LLMs is that they allow organizations to retain control over their data, thereby mitigating risks associated with external data processing and knowledge leakage [12]. Private deployments support compliance with regulatory frameworks such as the EU's General Data Protection Regulation (GDPR) [13], the EU AI Act, the United States' Executive Order 14110, Canada's AI and Data Act, China's Interim Measures for Generative AI Services, and Brazil's AI Bill [6]. Further risks associated with unauthorized external AI tool usage include cyber security vulnerabilities and compliance breaches [4].

Second, enterprise LLMs enable technical customization and domain adaptation. Models can be fine-tuned or augmented with internal datasets using retrieval augmented generation (RAG), allowing them to understand industry-specific terminology, workflows, and compliance requirements. This significantly improves performance in use cases such as knowledge management, workflow automation, customer support, and decision support systems [11].

Third, organizations benefit from improved integration with existing IS infrastructures and information systems. Many enterprises operate complex landscapes of legacy databases, enterprise resource planning systems, and fragmented data sources. LLMs can function as cognitive intermediaries or middleware, facilitating interaction across heterogeneous systems and improving interoperability while reducing configuration efforts [14]. LLMs can also be deployed in customer-facing roles, for example as customer support chatbots that respond to requests quickly, increasing customer satisfaction while reducing personnel costs. [9, 15].

Finally, governance and organizational control represent a major motivation for internal deployments. Enterprise LLM solutions allow organizations to manage model updates, maintain audit trails, implement risk management strategies, and apply ethical safeguards [16, 17].

3. Methodology

This study adopts a qualitative research design by investigating an exemplary case study of ITC, which is a pseudonymized leading information technology (IT) service provider headquartered in Germany. ITC has developed and deployed itcGPT, a GenAI application designed for internal use.

We conducted semi-structured interviews that followed a guideline derived from the research questions and the existing literature on enterprise AI and IT governance. The transcribed interviews are

available in the online repository. To complement the interview data and enhance construct validity, we analyzed internal documents. These documents provided contextual information about formal structures, decision-making processes, and technical designs. We conducted the interviews with experts who were involved in the conception, governance, and implementation of the enterprise LLM. Specifically, we interviewed the Data Protection Officer and Auditor (I1), the Head of the Competence Center Microsoft (I2), an Information Security Specialist (I3), the IT Demand Manager (I4), the Lead Enterprise Architect (I5), the Senior Manager IT Operations (I6), and the Team Lead IT Management (I7) in March 2025.

We analyzed the interview transcripts using qualitative content analysis, following the qualitative data analysis methodology by Mayring [18], in which we combined deductive and inductive elements. We first derived a three-part main categorization distinguishing between technical, organizational, and ethical requirements based on the research questions. Within these three overarching categories, subcategories were inductively developed. The combination of interviews and document analysis, together with a structured yet flexible coding procedure, ensured a robust empirical foundation for identifying and operationalizing the requirements for integrating enterprise LLMs into existing IS infrastructures.

4. Results

To answer our research questions, our analysis revealed five central requirements relevant for integrating enterprise LLMs into existing IS infrastructures. Although safety was not identified as a separate requirement category, it emerges as a cross-cutting property supported by the identified requirements. In particular, data quality and fairness, transparency, responsibility, and data privacy and security contribute to preventing harmful consequences of enterprise LLM use. We describe these requirements and outline corresponding operationalization measures, which are structured along the dimensions of the TOE framework. Figure 1 provides an overview of the key operationalization measures derived from our analysis. Structured along the technical, organizational, and ethical dimensions of the TOE framework, it illustrates how enterprise LLM integration can be supported in practice.

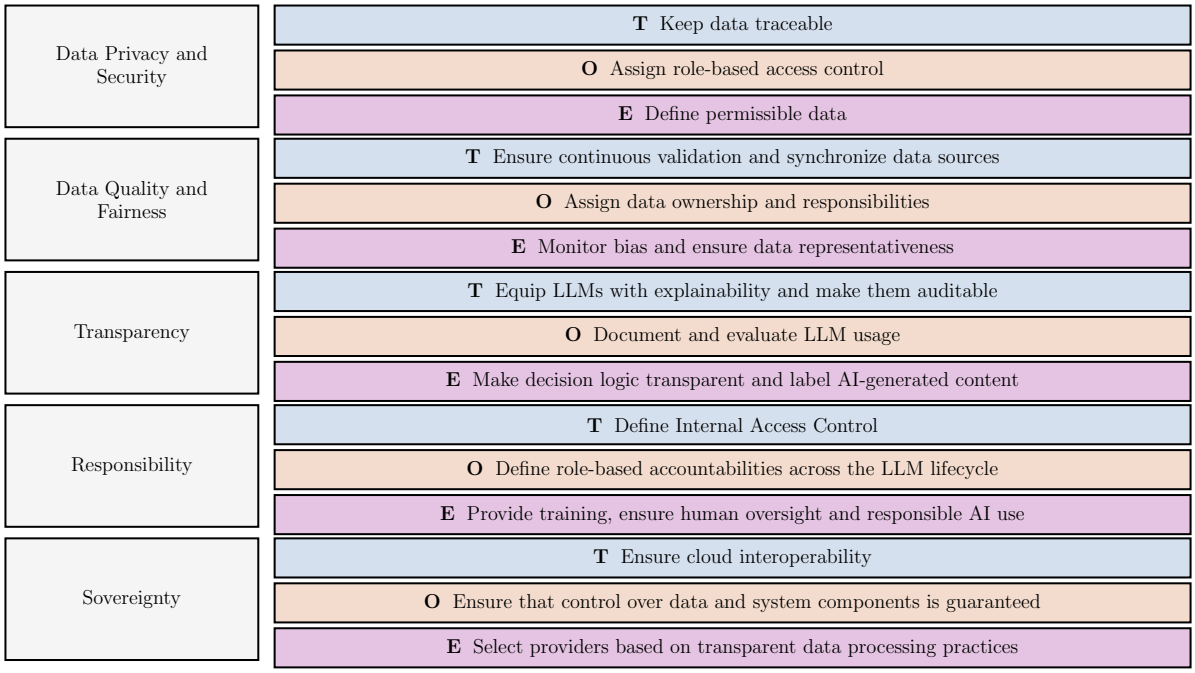

Figure 1: Overview of key technical (T), organizational (O), and ethical (E) operationalization measures for enterprise LLM integration.

Data Privacy and Security A central requirement for enterprise LLM integration is maintaining continuous control over data flows, storage locations, and processing environments. Sensitive corporate and personal data must remain within controlled infrastructures and defined jurisdictions. Compliance with data protection regulations and internal policies is essential (I1; I3; I6; I7). Responsible handling of entrusted data is not only a legal requirement but also necessary to maintain trust among employees, customers, and partners. In this sense, security measures also contribute to safety, but they do not exhaust it.

Operationalization: The use of secure data transfer protocols like TLS is advised. Data processing must remain purpose-bound, traceable, and compliant. Data governance should be embedded into the architecture through data catalogs, ownership rules, classification schemes, and monitoring of sensitive data flows (I7). Encryption, logical separation, secure transport standards, zero-trust principles, and continuous monitoring are essential.

Organizations must embed LLM use cases into formal governance workflows. Every new LLM application should pass a structured approval process before productive deployment. Role-based access control must be implemented via centralized identity management systems, with clearly assigned data owners and periodic access reviews. In addition, recurring compliance checks must be established to reassess data processing practices and ensure continued alignment with regulatory and internal policies. Responsibilities for system operation, data stewardship, and incident handling must be formally documented.

Organizations must prohibit the processing of private content within enterprise LLM systems and define policies specifying which internal data categories may be accessed or retrieved. Positive and negative lists of permissible data sources and document types must be defined. Wherever personal data is processed within enterprise repositories, pseudonymization or anonymization mechanisms must be embedded directly into data flows and retrieval processes (I1).

Data Quality and Fairness Enterprise LLM performance depends on the quality, completeness, and timeliness of data. This is particularly important when models retrieve information from internal documents, databases, or knowledge repositories. High data quality is also necessary to mitigate bias and unfair outcomes. Biases embedded in training data or internal knowledge repositories may be reproduced in model outputs and influence decision-support processes.

Operationalization: Organizations must ensure that the data is current, complete, and consistent (I7). Continuous validation, synchronized data sources, and ongoing quality assurance are necessary.

Organizations must implement formal data governance structures with clearly assigned data ownership and stewardship roles (I7). Structured documentation standards must be defined, and internal knowledge repositories should be periodically reviewed (I6). In addition, data contribution practices must be institutionalized. Employees should follow defined guidelines for documenting and structuring information to ensure machine-readability and contextual clarity. Regular quality checks and feedback mechanisms should be established to evaluate whether LLM outputs reflect reliable internal data (I7).

Where feasible, self-trained or internally fine-tuned models should be prioritized (I1). If externally pre-trained models are deployed, systematic bias assessments prior to and during enterprise use are mandatory. Measurable bias indicators and acceptable threshold values must be defined. Independent oversight mechanisms should be established to critically assess fairness-related risks.

Transparency Enterprise LLM integration requires transparency regarding system behavior, data sources, and organizational usage. In RAG-based settings, transparency depends on providing visibility into retrieved documents and metadata to prevent misinterpretation of internal content (I2). Transparency is essential to maintain trust and enable informed evaluation of AI-generated content. It also supports safety by enabling users and organizations to identify uncertain, incorrect, or potentially harmful outputs before they affect organizational processes.

Operationalization: Technical transparency requires complete logging, versioning, and documentation of all system-relevant processes. In enterprise LLM settings, this includes model changes, parameter settings, connected data sources, retrieval configurations, and output behavior. These records must remain audit-proof, regularly reviewed, and integrated into governance structures. In parallel, explainability must be ensured for both the model and its outputs, especially in RAG systems where retrieved

content and metadata shape responses. Explanations should therefore make the data basis, influencing factors, and uncertainties understandable for different user groups (I2; I5).

Organizations must ensure that every LLM-related application is registered within centralized governance instruments. New LLM use cases should only be deployed after being documented with defined ownership. Enterprise architecture tools must be continuously updated to reflect the current system landscape and to prevent redundant or uncontrolled deployments. In addition, license allocation and usage monitoring mechanisms must be implemented to provide insights into system adoption and cost development (I6). Access should be centrally managed and usage patterns regularly evaluated to identify uncontrolled proliferation or shadow usage (I2; I5; I6). Clear communication of approved tools and usage boundaries is required to prevent unauthorized reliance on public LLM solutions.

Users must be explicitly informed about how the enterprise LLM functions, which data sources are connected, and how retrieval mechanisms influence outputs (I2). AI-generated content must be clearly labeled. Furthermore, organizations must define exclusion criteria for high-risk enterprise use cases, such as legally sensitive, compliance-critical, or high-impact decision scenarios. (I3).

Responsibility Enterprise LLM integration requires clearly defined accountability for system deployment, operation, and oversight. From a safety perspective, such accountability is essential to ensure that LLM outputs do not directly translate into harmful decisions or actions without appropriate review.

Operationalization: Responsibility must be operationalized through strict access governance based on need-to-know and least-privilege principles. Enterprise LLM systems should enforce differentiated user roles consistently across applications, retrieval layers, data repositories, and interfaces. This should be supported by centralized identity services, continuous logging of security-relevant activities, and regular reviews of access rights. Authentication must follow modern security standards, including multi-factor authentication, cryptographic safeguards, and separate administrative and operational access paths, with privileged accounts receiving additional protection (I3; I4; I6).

Organizations must assign role-based accountability across the entire LLM lifecycle. The roles must be embedded within existing governance processes. (I6). In addition, human oversight must be structurally integrated into LLM-supported workflows (I1). Critical outputs must require validation, such as approval procedures or dual-control mechanisms. Responsibilities for output verification, incident handling, and ongoing monitoring must be explicitly defined.

In automated or semi-automated workflows, human intervention must remain possible. AI-generated outputs should be reviewed through structured procedures to assess not only accuracy but also ethical, communicative, and contextual appropriateness (I1). Enterprise LLMs should support, not replace, human expertise (I4). Decisions with major financial, legal, or organizational consequences must not rely solely on automated outputs. Access to enterprise LLMs should be equal across departments and hierarchy levels. Systems must be understandable and usable for diverse groups, with training for responsible use.

Sovereignty Maintaining sovereignty helps preserve long-term organizational autonomy and reduces exposure to external legal, economic, or political influences (I6). Strong vendor dependencies reduce portability and limit future migration options, making early architectural consideration essential (I5; I6).

Operationalization: Technical sovereignty depends on interoperability, modularity, and scalable deployment (I4). Enterprise LLM systems should support standardized interfaces, structured data formats, and bidirectional communication with existing systems. Architectures should remain portable across on-premises and private cloud environments through modular separation of storage, retrieval, and inference components, and clearly defined integration points.

Each LLM application should undergo an evaluation regarding hosting location, dependency on specific cloud providers, and portability constraints before approval (I6). Architectural principles must include criteria for avoiding unnecessary vendor lock-in and ensuring compatibility with existing IS infrastructure. Furthermore, operating models must be documented and periodically reviewed to verify continued control over data and system components. Organizations should ensure that critical data remains within approved infrastructure boundaries and that alternative deployment options remain technically feasible.

Organizations must treat provider selection as an ethical and strategic decision. European, providers

should be preferred because they offer higher transparency regarding data processing, storage, and model governance. Selecting transparent and accountable providers strengthens traceability, supports regulatory alignment, and reinforces long-term organizational control.

5. Discussion and Outlook

This study examined the integration of enterprise LLMs into existing IS infrastructures by identifying technical, organizational, and ethical requirements (RQ1) and deriving corresponding operationalization measures (RQ2). It contributes to research and practice by showing that enterprise LLM integration is not a purely technical implementation challenge but a socio-technical transformation spanning technological, organizational, and ethical dimensions. The derived requirements and operationalization measures provide actionable guidance for organizations implementing enterprise LLMs.

Our findings further highlight the importance of explainability, transparency, and safety properties. In particular, data privacy and security as well as transparency emerged as critical requirements. Ensuring transparency requires not only documentation and traceability of system behavior but also explainability of LLM outputs and their underlying data sources. However, because generated reasoning traces and post-hoc explanations may appear plausible without faithfully representing the model's actual decision process, organizations should complement explainability with governance mechanisms such as systematic validation, continuous monitoring, and human oversight [19]. Without mechanisms that allow users and organizations to understand how outputs are generated, it becomes difficult to assess reliability, detect risks, or ensure compliance.

Importantly, contrary to previous publications, we suggest that deploying an enterprise LLM does not relieve organizations of concerns such as data protection, auditability, or compliance [11]. Responsibilities that might otherwise be partially outsourced to external AI providers are shifted to the implementing organization itself. Operating privately deployed enterprise LLMs therefore increases the need for internal governance structures, monitoring mechanisms, ethical considerations, and clearly defined accountability. However, if organizations systematically implement the identified requirements and operationalization measures, they can benefit from increased control, transparency, and sovereignty over their data and AI-driven processes. Notably, participants rarely considered on-premises hosting of language models as an alternative to a private cloud option that could reduce provider dependencies and provide organizations with greater control over sensitive data and system operations [2].

Although the dimensions of the TOE framework strongly overlap in practice, our findings reveal that tensions may arise between these dimensions. One central tension concerns responsibility and accountability. Organizations are typically required to appoint responsible individuals in case system failures occur. However, when decisions are automatically generated or supported by AI systems, failures often result from complex interactions between training data, model architecture, retrieval mechanisms, user prompts, and system integrations. This distributed causality complicates the clear attribution of responsibility. To mitigate this challenge, maintaining a "human-in-the-loop" becomes critical, particularly for high-impact or compliance-relevant use cases. However, human oversight may increase processing time, personnel costs, and organizational bottlenecks, thereby reducing expected efficiency gains. Organizations should therefore adopt a risk-based oversight model: low-risk applications may rely on automated monitoring or ex-post review, whereas high-risk outputs should require mandatory human validation, approval workflows, or dual-control mechanisms. This helps balance efficiency and safety by concentrating human oversight where potential harms and legal consequences are greatest.

Some limitations must be acknowledged. First, the study is based on a single-case design. While this limits statistical generalizability, the consultancy context of the investigated organization strengthens the transferability of the findings. The interviewees had experience with software implementation projects across different clients, industries, and organizational settings, allowing them to reflect on recurring challenges beyond a single internal use case. Furthermore, the derived requirements and operationalization measures are formulated in a general manner and are not specific to the investigated firm. They may therefore offer guidance for other organizations implementing enterprise LLMs, without

claiming exhaustiveness or universal applicability. Nevertheless, the empirical insights remain rooted in one organizational context and may be influenced by subjective interpretations. The derived framework therefore remains exploratory rather than explanatory. Moreover, the data collection focused on executives and IT architects and did not include end users, potentially overlooking adoption barriers, shadow AI practices, and workflow frictions that influence the actual integration and use of AI within organizations.

A further limitation is that the interviews primarily surfaced established IT security measures, while LLM-specific risks such as prompt injection, hallucinations, sensitive data leakage, or model inversion attacks were not explicitly emphasized. Future work should therefore examine how such safeguards can complement generic security requirements in enterprise LLM implementations.

Moreover, this study does not address the prior strategic question of whether implementing an enterprise LLM is advisable in the first place. Organizations should carefully evaluate whether an LLM addresses a clearly defined organizational need and establish appropriate success criteria before implementation; however, our study focuses on cases in which the decision to integrate an enterprise LLM has already been made.

Additional perspectives, such as economic cost considerations or ecological sustainability implications of enterprise LLM deployment, were beyond the scope of this study but represent promising directions for further research. As enterprise AI ecosystems mature, future work may refine, extend, and empirically test the proposed framework in light of these additional dimensions. Additionally, future research should examine how trust challenges and corresponding governance requirements differ between RAG-fine-tuned LLMs and agentic architectures that integrate external tools and databases through interfaces and may pose a future alternative to fine-tuning.

Acknowledgments

This research was conducted within the Erasmus+ Cooperation Partnership FAIR PLAI (Fostering Awareness of Integrity and Responsibility in Problem-based Learning of AI), which promotes a holistic perspective on AI with a particular focus on ethical considerations.

Declaration on Generative AI

During the preparation of this work, the authors used GPT 5.5 in order to: Improve writing style. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] S. Feuerriegel, J. Hartmann, C. Janiesch, P. Zschech, Generative ai, Business & Information Systems Engineering 66 (2024) 111–126.

[2] S. Lins, K. D. Pandl, H. Teigeler, S. Thiebes, C. Bayer, A. Sunyaev, Artificial Intelligence as a Service, Business & Information Systems Engineering 63 (2021) 441–456.

[3] M. Markic, S. Webster, A. van der Staay, F. N. Bäßmann, C. Janiesch, J. Poeppelbuss, Advancing AI Adoption in SMEs: A Framework for Nascent AI Strategy Development, in: ICIS 2025 Proceedings, 2025.

[4] D. Puthal, A. K. Mishra, S. P. Mohanty, A. Longo, C. Y. Yeun, Shadow ai: Cyber security implications, opportunities and challenges in the unseen frontier, SN Computer Science 6 (2025) 405.

[5] M. Silic, D. Silic, K. Kind-Trüller, From shadow it to shadow ai–threats, risks and opportunities for organizations, Strategic Change (2025).

[6] S. Alanoca, S. Gur-Arieh, T. Zick, K. Klyman, Comparing apples to oranges: A taxonomy for navigating the global landscape of ai regulation, Association for Computing Machinery, New York, NY, USA, 2025.

[7] L. Gkinko, A. Elbanna, The appropriation of conversational ai in the workplace: A taxonomy of ai chatbot users, International Journal of Information Management 69 (2023) 102568.
[8] T. Baldazzi, L. Bellomarini, S. Ceri, A. Colombo, A. Gentili, E. Sallinger, Fine-tuning large enterprise language models via ontological reasoning, in: International Joint Conference on Rules and Reasoning, Springer, 2023, pp. 86–94.
[9] A. van der Staay, M. Markic, P. Stahmann, M. Nebel, C. Knickrehm, J. Poeppelbuss, C. Janiesch, AI-as-a-Service Mediation: How Ordinary Companies Become AIaaS-powered Service Providers, in: ECIS 2025 Proceedings, 2025.
[10] M. Bosch-Rekveldt, Y. Jongkind, H. Mooi, H. Bakker, A. Verbraeck, Grasping project complexity in large engineering projects: The toe (technical, organizational and environmental) framework, International journal of project Management 29 (2011) 728–739.
[11] D. E. O'leary, D. E. O'leary, The rise and design of enterprise large language models, IEEE Intelligent Systems 39 (2024) 60–63.
[12] D. E. O'Leary, Enterprise large language models: Knowledge characteristics, risks, and organizational activities, Intelligent systems in accounting, finance and management 30 (2023) 113–119.
[13] F. Mahr, G. Angeli, T. Sindel, K. Schmidt, J. Franke, A reference architecture for deploying large language model applications in industrial environments, 2024 IEEE 30th International Symposium for Design and Technology in Electronic Packaging (SIITME) (2024) 19–23.
[14] N. V. K. A. Vedanbhatla, Llms as ai middleware: Unifying disparate systems in manufacturing it landscapes, Journal of Computer Science and Technology Studies (2025).
[15] R. Urbani, C. Ferreira, J. Lam, Managerial framework for evaluating ai chatbot integration: Bridging organizational readiness and technological challenges, Business Horizons 67 (2024) 595–606.
[16] T. R. McIntosh, T. Susnjak, T. Liu, P. Watters, D. Xu, D. Liu, R. Nowrozy, M. N. Halgamuge, From cobit to iso 42001: Evaluating cybersecurity frameworks for opportunities, risks, and regulatory compliance in commercializing large language models, Computers & Security 144 (2024) 103964.
[17] J. Hu, D. Wang, Z. Wang, X. Pang, H. Xu, J. Ren, K. Ren, Federated large language model: Solutions, challenges and future directions, IEEE Wireless Communications 32 (2025) 82–89.
[18] P. Mayring, T. Fenzl, Qualitative inhaltsanalyse, in: Handbuch methoden der empirischen sozialforschung, Springer, 2019, pp. 633–648.
[19] M. Turpin, J. Michael, E. Perez, S. Bowman, Language models don't always say what they think: Unfaithful explanations in chain-of-thought prompting, Advances in Neural Information Processing Systems 36 (2023) 74952–74965.