Diener, Moritz; Kaps, Simon; Spitzer, Philipp; Hirt, Robin; Vössing, Michael; Satzger, Gerhard (2026): Rethinking Knowledge Work: Designing LLM-based Systems for Complexity Management. *Proceedings of the 34th European Conference on Information Systems (ECIS 2026).*

Acknowledgements: This work has been supported by the German Federal Ministry of Education and Research through the research project “bi.smart” (grant no. 02J19B041).

Karlsruhe Institute of Technology (KIT)
Kaiserstraße 12
76131 Karlsruhe
Germany

prenode GmbH
Rintheimer Straße 23
76131 Karlsruhe
Germany

RETHINKING KNOWLEDGE WORK: DESIGNING LLM-BASED SYSTEMS FOR COMPLEXITY MANAGEMENT

Completed Research Paper

Moritz Diener, Karlsruhe Institute of Technology, Karlsruhe, Germany, moritz.diener@kit.edu

Simon Kaps, prenode GmbH, Karlsruhe, Germany, simon.kaps@prenode.de

Philipp Spitzer, Karlsruhe Institute of Technology, Karlsruhe, Germany, philipp.spitzer@kit.edu

Robin Hirt, prenode GmbH, Karlsruhe, Germany, robin.hirt@prenode.de

Michael Vössing, Karlsruhe Institute of Technology, Karlsruhe, Germany, michael.voessing@kit.edu

Gerhard Satzger, Karlsruhe Institute of Technology, Karlsruhe, Germany, gerhard.satzger@kit.edu

Abstract

Knowledge workers increasingly face complexity in managing heterogeneous documents, strict compliance requirements, and cross-functional processes, which disproportionately burden small and medium-sized enterprises (SMEs). Particularly in public procurement, tender analysis involves processing heterogeneous documents, complying with requirements, and coordinating across functions to inform bid decisions. Using design science research with an SME, we investigate how to manage complexity in document-intensive knowledge work—exemplified by public tender analysis—across document, compliance, and process dimensions, deriving four design requirements and two design principles. We instantiate these in a Large Language Model (LLM)-based artifact that supports document analysis. Evaluated on real tenders, the artifact reduces initial screening time while maintaining high retrieval accuracy. Our findings demonstrate that knowledge work can be advanced by LLMs, reducing complexity and transforming the nature of human work—shifting the focus from reading and extracting information to orchestrating and verifying LLM-generated outputs.

Keywords: large language models, generative AI, complexity management, knowledge work.

1 Introduction

Organizations increasingly face complexity in knowledge-intensive processes requiring extensive documentation, strict compliance, and cross-functional coordination (Eppler & Mengis, 2008; Okhuysen & Bechky, 2009; Voglhofer & Rinderle-Ma, 2020). Knowledge workers must navigate heterogeneous documents, extract critical information under time pressure, meet formal obligations where even minor oversights can have severe consequences, and coordinate across departments. Digitalization has increased the volume, velocity, and heterogeneity of information that must be processed (Faraj et al., 2018; Vuori et al., 2019), often exacerbating rather than reducing complexity for knowledge workers (Wunderlich & Fischer, 2022). Public procurement exemplifies these challenges: bidders must analyze tender documents spanning hundreds of pages in non-standardized formats, comply with strict requirements where a single missing certificate results in immediate disqualification, and coordinate across sales, legal, and technical teams under tight deadlines. Recent advances in Generative Artificial Intelligence (GenAI)—particularly Large Language Models (LLMs)—offer new possibilities for addressing such complexity through document analysis, summarization, and information retrieval

(Söllner et al., 2025). However, prescriptive design knowledge on how to design LLM-based systems that effectively address complexity remains limited.

Complexity in document-intensive, compliance-critical knowledge work manifests across multiple dimensions. First, complexity can be viewed as a mismatch between information demands and human processing capacity, which impacts decision quality and work performance through information overload (Eppler & Mengis, 2008; Roetzel, 2019). This emerges when knowledge workers must screen large amounts of heterogeneous information under time pressure, particularly in non-standardized document formats. Beyond information overload, research on coordination in organizations highlights that interdependent tasks distributed across roles and departments require substantial organizing work—such as creating shared understandings, negotiating responsibilities, and aligning timelines (Okhuysen & Bechky, 2009), representing a second dimension: process complexity. Additionally, in compliance-critical contexts, detailed rules, auditability requirements, and evolving regulations introduce a third dimension—compliance complexity—that must be embedded in processes and information systems (IS) (Voglhofer & Rinderle-Ma, 2020). These three dimensions—document complexity (information overload), process complexity (coordination demands), and compliance complexity (regulatory requirements)—interact rather than exist as separate challenges in document-intensive knowledge work.

LLMs offer capabilities to address these complexities and have been viewed as a new type of component in socio-technical systems that can automate cognitive tasks for knowledge workers such as drafting, summarizing, classifying, and coding (Feuerriegel et al., 2024). Studies have highlighted that LLMs can assist with search, retrieval, and knowledge sharing by transforming unstructured documents into more organized, queryable formats, thereby reshaping how work is conducted (Jarrahi et al., 2023; Kirchner et al., 2025; Storey, 2025). Empirical work shows that LLMs can augment key tasks such as screening, information retrieval, summarization, and synthesis, but also introduce new risks, including hallucinations, opaque reasoning, and overreliance on LLM-generated outputs (Benbya et al., 2024). To mitigate this, prior work proposes that LLM-based systems need to be designed to augment rather than replace knowledge work, preserve human agency, and support accountable decision-making through appropriate control and transparency mechanisms (Fügener et al., 2021; Krakowski, 2025; Shneiderman, 2020). However, existing research primarily focuses on generic productivity support or conceptual discussions of LLMs in organizations. Prescriptive design knowledge on how LLM-based IS can address document, compliance, and process complexity in high-stakes knowledge work remains limited. Thus, we set out the following research question:

How can LLM-based IS be designed to support knowledge workers in managing complexity across document, compliance, and process dimensions?

To address this research question, we apply a design science research (DSR) approach in the context of bidder-side public tender analysis where document-intensive, compliance-critical knowledge work and all three complexity dimensions converge: bidders process heterogeneous documents spanning hundreds of pages under tight deadlines, face strict formal requirements where errors lead to disqualification, and coordinate across sales, legal, and technical departments. Moreover, SMEs face these challenges with limited resources, making effective complexity management essential for their participation in public procurement. Through collaboration with a German SME that operates data centres and regularly participates in public tenders, we conducted interviews with domain experts to identify challenges and requirements, derived design principles (DPs) grounded in kernel theories, instantiated these principles in an LLM-based artifact, and evaluated the artifact through quantitative and qualitative methods.

Overall, the contributions of this paper are threefold: First, we identify and operationalize three dimensions of complexity in document-intensive, compliance-critical knowledge work—document, compliance, and process complexity—grounded in the procurement context, and derive four design requirements (DRs). Second, we propose and instantiate design knowledge in the form of two DPs for LLM-based systems that support knowledge workers in managing these complexities: the LLM-based artifact demonstrates high information retrieval accuracy and reduction in initial screening time, complemented by a qualitative think-aloud study that evidences reductions in document, compliance, and process complexity. Third, we observe an emergent shift in work pattern: knowledge workers (1) orchestrate LLM-based systems to retrieve key information into structured overviews, checklists, and timelines, and (2) verify and correct outputs using source links and expert knowledge, suggesting how LLM-based systems shift knowledge work from information extraction to output verification.

2 Related Work

2.1 Complexity Management in Document-Intensive Knowledge Work

Document-intensive knowledge work with strict compliance requirements, and cross-functional coordination expose knowledge workers to multiple forms of complexity. Information overload arises when heterogeneous documents must be screened and compared under time pressure, while coordination complexity stems from the need to align distributed experts and roles (Okhuysen & Bechky, 2009). Compliance complexity describes the challenge of executing work under detailed rules where minor errors can have severe consequences (Voglhofer & Rinderle-Ma, 2020). Digitalization does not simply reduce these complexities, but can transform knowledge work and how organizations are structured (Faraj et al., 2018; Wunderlich & Fischer, 2022). GenAI and LLMs are increasingly discussed as socio-technical systems for managing such complexity (Harder Fischer et al., 2024). LLMs provide natural-language interfaces for summarization, extraction, and question answering over documents, redistributing cognitive effort from manual reading toward interpretation and decision-making (Feuerriegel et al., 2024). Retrieval-augmented generation (RAG) grounds outputs in organizational data such as internal documentation, contracts or policies, and emerging multi-agent systems orchestrate multiple LLM-based agents across workflows (Klesel & Wittmann, 2025). From a socio-technical view, human–AI collaboration and responsible AI research stress that such systems must preserve human oversight, accountability, and control in high-stakes, compliance-critical contexts (Amershi et al., 2019; Shneiderman, 2020) and that LLM outputs must remain explainable and reviewable so that experts can verify, correct, and justify decisions (Feuerriegel et al., 2024; Hassani, 2024; Söllner et al., 2025).

2.2 Kernel Theories for Complexity Management

To situate our work in existing theory, we draw on research that explains how IS shape document analysis, information processing, and complexity management for knowledge workers.

Information overload. Information overload research views complexity as a mismatch between information demands and human processing capacities: when the volume, heterogeneity, and fragmentation of information exceed what individuals can handle, attention and decision quality deteriorate (Eppler & Mengis, 2008; Roetzel, 2019). In response, countermeasures such as filtering, summarization, structuring, and visualization are suggested to keep information loads within manageable bounds. This perspective frames complexity management as designing systems that externalize manual reading, structuring, and information processing, transforming long, fragmented documents into concise, structured, and traceable summaries for knowledge workers.

Socio-technical systems. Socio-technical systems (STS) theory views organizations as joint systems of social actors and technical artifacts that must be co-designed so that tasks, responsibilities, and organizational structures meaningfully align (Mumford, 2006; Trist, 1981). STS theory shows that misalignment increases coordination complexity and ambiguity. In contrast, well-designed systems help knowledge workers manage organizational and task complexity by clarifying information and aligning the knowledge workers' needs with the capabilities of technologies. This implies that complexity management cannot be achieved solely by technology, but through STS that deliberately redistributes information retrieval and coordination between humans and systems, rather than simply shifting or amplifying complexity.

Task–technology fit. Task-technology fit (TTF) theory argues that IS support knowledge workers when their functionalities align with task requirements and user capabilities, rather than focusing solely on technology (Goodhue & Thompson, 1995). When this fit is low, systems create additional effort and confusion for knowledge workers, thereby increasing complexity for them. For complexity management, TTF highlights that systems must provide outputs that align with core subtasks of knowledge workers, thereby reducing complexity rather than adding new layers of cognitive and coordination effort.

3 Methodology

3.1 Research Approach

Our study follows the DSR process outlined by Kuechler & Vaishnavi (2008), which consists of five stages: awareness, suggestion, development, evaluation, and conclusion. The **relevance** is ensured through collaboration with a German SME active in bidder-side public procurement. We selected this SME due to its frequent participation in tenders and its difficulties in handling document-intensive, compliance-critical knowledge work. The **rigor cycle** draws on prior literature and kernel theories. We derived design knowledge in the form of DPs, instantiated these in an artifact, and evaluated through a quantitative evaluation and a think-aloud study.

Awareness. We jointly built problem awareness with a German SME that operates data centres and regularly bids in public procurement. We conducted a two-hour workshop followed by semi-structured stakeholder interviews (n = 7). A purposeful sampling of experts for interviews was employed to ensure the representation of key expert roles in tender analysis workflows. Table 1 displays interview participants by roles, research phase, and interview durations. Following a two-stage operationalization process (Kaiser, 2021; Kallio et al., 2016), a semi-structured interview guideline was developed to cover existing complexities and challenges, as well as the desired requirements for a solution. The interviews were recorded, transcribed, and a systematic qualitative content analysis (Mayring, 2000) with inductive coding (Gioia et al., 2013) was employed. Interview findings were cross-verified with the company's internal documents provided by the participants.

Role	Awareness: Pre-Study	Awareness: Interviews	Evaluation: Think-Aloud
Contract Manager (1)		0:58h	0:23h
Contract Manager (2)		0:37h	
Public Account Manager		0:49h	
Sales Manager	2:00h		
Sales Representative	2:00h	1:02h	0:33h
Senior Software Consultant		0:47h	
Technical Sales Manager		0:50h	0:47h
Technical Sales Consultant		1:03h	
Head of Marketing			0:29h

Table 1. Overview of interviewees categorized by role, research phase, and duration.

Suggestion and Development. Based on the codes developed from the interviews, we derived four DRs for IS and subsequently formulated two DPs using kernel theories (Gregor et al., 2020). The DPs were instantiated in an artifact that addresses the problem of document, compliance, and process complexities for knowledge workers.

Evaluation. Following the Framework for Evaluation in DSR (Venable et al., 2016), we employed the *Technical Risk & Efficacy* evaluation strategy. First, we conducted a quantitative evaluation of our artifact using real-world tender documents to assess the reduction in screening time and information retrieval accuracy. In consultation with the tender expert of the SME, five target information items as structured text outputs and two binary yes/no questions were defined as retrieval objectives. These objectives represent the generic topics of interest in a tender, including, for example, the content of the tender and the requirements to participate. Each document was manually annotated to establish the ground truth. The artifact's outputs were compared to this ground truth, and the accuracy of information retrieval was measured. To evaluate screening time, experts manually retrieved the information objectives from the documents, and this duration was recorded and compared to the processing time of the artifact, from uploading the documents to displaying the information. Second, structured think-aloud sessions (Ericsson & Simon, 1993) were conducted (n = 5), in which participants completed realistic tender analysis tasks while verbally articulating their thought processes. During these sessions, the perceived reduction in complexity was assessed, supplemented by post-session interviews that were transcribed and deductively coded (Mayring, 2000).

Conclusion. We synthesize the findings from the evaluation and document the design knowledge for IS that support humans in reducing complexity when managing heterogeneous documents, formal compliance requirements, and cross-functional processes.

3.2 Application Context

Tendering is a formal, competitive process where a contracting authority, such as government agencies and public institutions, invites bids for specific works, supplies, or services from private companies (Flynn, 2025). Through open competition, tendering ensures equal treatment of companies, transparency in public spending, and value for money. Public invitations to tender account for approximately 15% of global GDP (Fazekas & Blum, 2021). In Germany alone, a total value of €123.5 billion was distributed through 195,493 procurement procedures in 2023 (Statistisches Bundesamt, 2025). Yet, it remains a time-consuming, error-prone, and costly process for companies to identify opportunities and decide whether to bid on a call for tenders since creating a tender proposal requires extensive collaboration between sales, legal, financial, and technical departments (Flynn et al., 2015). Additionally, SMEs often face challenges in participation due to limited resources and market access barriers (Flynn et al., 2015). Despite advances in digitalization, tender analysis remains document-intensive, inefficient, and procedurally complex (Oussaleh Taoufik & Azmani, 2023). Procurement research has mainly concentrated on the buyer or evaluator side of the process, including the automation of supplier assessment, bid evaluation, and sourcing decisions (Loader, 2015; Vaidya et al., 2006). However, gaps remain on the bidder side, especially in knowledge work supported by LLM-based systems.

Interviews with stakeholders from the partner SME, including sales, contract management, and technical roles, reveal a five-step workflow that bidding companies follow to identify tender opportunities, analyze tender documents, make a bid decision, and write a bid proposal. First, teams continuously scan portals like Tenders Electronic Daily and circulate promising calls for tender within the company, while performing an initial evaluation against the company's capabilities and strategy. Because the calls for tender are numerous and their attachments are lengthy and heterogeneous, early screening requires considerable time and effort. Effective opportunity screening, thus, is a success factor for participation in tenders (Loader, 2015). Second, a thorough preliminary check informs the bid/no-bid decision by weighing factors such as eligibility, capacity, required certifications, timelines, costs, expected competition, and win probability. Interviewees reported spending dozens of hours manually reading multiple PDFs and annexes, as well as frequently asking for clarifications due to ambiguous requirements. These are symptoms of document, compliance and process complexities that lead to information overload and fragmentation, which are known to impair decision quality since effective internal coordination during this phase can have a decisive impact on tender outcomes (Eppler & Mengis, 2008; Flynn & Davis, 2017). Third, if the decision is to proceed and to bid, a cross-functional kickoff meeting allocates work across sales, product/engineering, legal, finance, and contract management. Coordinating these efforts is essential but challenging in practice because of calendar conflicts and the perception among professionals in our interviews that tender preparation is burdensome and complex extra work. This issue is consistent with evidence on collaboration barriers and the importance of coordination capability (Flynn, 2017; McCue & Roman, 2012). Fourth, during the bid preparation process, teams craft technical and financial offers, collect relevant certificates, and tailor the content to meet the required specifications. Interviewees noted limited reuse due to the absence of a robust repository and inconsistent document structures. This is consistent with the role of standardization in reducing processing inefficiencies (Vaidya et al., 2006). Finally, bids are submitted under strict deadline and format requirements. Authorities then evaluate the bids, often selecting the lowest price or using the most economically advantageous tender method, which combines price, quality, and performance (Flynn & Davis, 2017).

4 Results

4.1 Awareness of Complexity in Knowledge Work

The systematic coding of the interviews revealed three dominant themes that summarize the fundamental challenges and complexities within the tender analysis process. Each theme covers multiple dimensions that collectively illustrate the complexity of public procurement activities.

Document Complexity. The interview analysis revealed three key aspects of document-related challenges. (1) Volume of Documentation: Tender documents frequently span hundreds of pages with numerous attachments, creating substantial information overload. As one Contract Manager noted, the sheer volume can easily span several hundred pages (“In public tenders, documents can easily be several hundred pages long.”), making it extraordinarily difficult to maintain an overview and identify critical information, comparable to finding a "needle in a haystack." (2) Non-Standardized Format: Tenders lack a uniform document structure, with each one exhibiting different layouts and organization. Critical details are often spread across multiple files and attachments rather than being centralized in one location, forcing suppliers to adjust their review approach for each new opportunity repeatedly. (3) Time Investment: The screening and preparation process demands considerable resource allocation. Interviewees reported “dozens of working hours“ for a thorough review, information gathering and final document verification, creating a substantial operational burden.

Compliance Complexity. The procedural requirements in public procurement and tender analysis present two related challenges. (1) Strict Formal Requirements: Public tenders enforce precise rules where minor deviations can result in immediate disqualification. As one expert pointed out, even a single missing certificate or a late submission can cause immediate disqualification (“[...] even a single piece of missing evidence can lead to exclusion. Formal errors, such as late uploading of documents, usually also lead to immediate exclusion“), creating risk throughout the process. (2) Unclear Requirements: Tender specifications frequently contain ambiguities that necessitate formal clarification through official portals. This procedural requirement introduces interruptions and delays as teams wait for responses from the contracting authority (“[...] have to wait until the tendering authority responds”) before proceeding with their work.

Process Complexity. The operational process within a company reveals multiple organizational challenges. (1) Cross-Department Coordination: Tender responses require extensive collaboration across sales, product, legal, finance, and contract management departments. While necessary, this cross-functional approach introduces substantial coordination overhead and potential communication gaps. (2) Internal Tender Fatigue: As noted by a Sales Representative, tender preparation is widely perceived as labor-intensive additional work, requiring management to actively motivate staff participation and engagement, which in turn impacts organizational efficiency. (3) Limited Knowledge Reuse: Organizations typically lack a centralized repository of past proposals or standard responses. Without a central tool that archives past requirements or automatically recognizes relevant content ("Currently, there is no central tool that fully archives all requirements ever submitted or operates with an automated recognition mechanism."), teams must start all over again for each new tender opportunity.

In summary, the interviews revealed that bidder-side tender analysis is a highly complex, knowledge-intensive process. This process involves voluminous and heterogeneous documents, strict (and sometimes ambiguous) compliance requirements, and cross-departmental coordination with limited knowledge reuse. These factors create a substantial cognitive and organizational burden.

4.2 Suggestion and Development of the Artifact

4.2.1 Design Requirements

Building on the three complexity dimensions identified in the awareness phase, we derived four DRs through inductive coding of the interviews. DRs denote context-specific requirements that the artifact must satisfy in order to support tender analysis.

DR1: Rapid Screening and Summarization of Tenders. The system should automatically analyze tender documents and create concise summaries that highlight key details such as the scope of work, eligibility criteria, certifications, and deadlines, reducing manual effort during screening. It should allow teams to assess tenders quickly without reviewing hundreds of pages. Interviewees emphasized this need ("If I could put all the documents into one system and get a summary, such as which certifications are required [...] and whether we can cover that in principle, then we would very quickly have an initial overview"). In line with previous research on complexity reduction, structured summaries that condense key information into different levels of detail can decrease cognitive load and help workers manage information overload (Eppler & Mengis, 2008).

DR2: Automated Deadline and Compliance Extraction. The artifact should automatically identify deadlines and compliance requirements in tender documents and present them as an actionable checklist and timeline. By externalizing obligations into structured reminders, such systems can help lessen human memory failures and lower the risk of overlooking critical requirements that cause immediate disqualification (Morrison et al., 2024). DR2 thus embeds regulatory awareness directly into the tender analysis, ensuring that procedural rules are systematically addressed while allowing knowledge workers to focus on the substantive quality and competitiveness of the bid.

DR3: Dual-Phase Analysis with Human-in-the-Loop. The artifact should incorporate two analysis phases. An initial rapid evaluation for quick bid decisions that filters unsuitable tenders quickly, followed by a detailed review under human supervision for bid proposal preparation. Importantly, outputs must remain subject to human oversight. Interviewees favored quick insights during initial screening but emphasized accuracy and control during a detailed analysis phase with human oversight. ("When it comes to pure prescreening, I prefer speed [...] At the beginning, I need fast information above all else [...] accuracy can then be ensured through human review [...]"). Kudina & van de Poel (2024) emphasize that humans must supervise LLM outputs for governance, legal reasons, and effectiveness, as LLMs might produce incorrect outputs and recommendations.

DR4: User-Friendly Interface and Workflow Integration. The artifact should provide an intuitive interface that fits seamlessly into existing tender workflows, with a simple graphical front-end for uploading documents, reviewing documents and LLM outputs, and posing follow-up queries. It must be usable by technical, legal, and sales staff without in-depth technical knowledge. One expert stressed that a command-line-only tool would not be sufficient ("A clear, intuitive interface would therefore be a great advantage. Command line inputs would be too complicated for most people."). This enhances both perceived ease of use and the likelihood of adoption in tender analysis (Davis & Granić, 2024).

4.2.2 Design Knowledge

Building on these DRs, we formulated two generalizable DPs using the framework of Gregor et al. (2020). The DPs express reusable prescriptive knowledge about how LLM-based systems can be designed to address document, compliance, and process complexity. Figure 1 summarizes the relationship between each DP, its DRs, complexity dimension and theoretical foundations.

DP1: LLM-based Document Analysis for Complexity Reduction. For designers of LLM-based IS that support knowledge workers, the aim is to reduce document complexity, enabling users to quickly get an overview of key information from large, heterogeneous tender documents (DR1, DR2, DR4). By reducing the complexity of large, heterogeneous documents through structured summaries, the system helps users manage information overload and supports task–technology fit between knowledge workers' screening tasks and system capabilities. This can be achieved through retrieval-based mechanisms such as RAG, which identify relevant passages across large collections of documents, synthesize them into task-oriented overviews, and retain links to the original text so that systems outputs are verifiable. This principle is based on research about information and cognitive overload, which shows that too much unstructured information harms human performance (Eppler & Mengis, 2008), and on TTF theory, which demonstrates that aligning system capabilities with users' tasks of complex tender analysis enhances individual performance and perceived usefulness (Goodhue & Thompson, 1995). Thus: employ LLM-based document analysis that automatically retrieves, condenses, and structures key information from documents into task-aligned summaries and overviews that are linked to their underlying source passages from the documents.

DP2: Human-based Verification and Human-AI Orchestration. For designers of LLM-based IS that support knowledge workers, the aim is to reduce process and compliance complexity by shifting the dominant work mode from reading and manually extracting information toward orchestrating and verifying LLM-generated outputs. Systems should present LLM-generated outputs of key information as checklists, timelines, and structured overviews (DR2, DR3, DR4). Orchestration refers to the human coordination and interaction with one or more LLM-based systems specialized for different subtasks, directing queries, deciding which system outputs to use for each task, and determining when additional follow-up or deeper analysis is required. Verification refers to the human evaluation of these systems' outputs by checking their faithfulness to linked source passages, assessing whether critical information is missing, and judging whether the interpretation is plausible in light of domain expertise before consequential decisions are made. This principle is grounded in STS theory, which advocates alignment of technical structures and human roles and emphasizes meaningful human participation and oversight (Kudina & van de Poel, 2024; Trist, 1981), as well as in human-centered and responsible AI perspectives, which stress that LLM-generated outputs must remain explainable and reviewable so that experts can verify, correct, and justify decisions (Feuerriegel et al., 2024; Hassani, 2024; Söllner et al., 2025). Thus: provide a dual-phase, source-grounded analysis in which knowledge workers (1) orchestrate LLM-based systems to retrieve key information into structured overviews, checklists, and timelines for initial screening and coordination, and (2) verify and correct outputs using source links and expert knowledge.

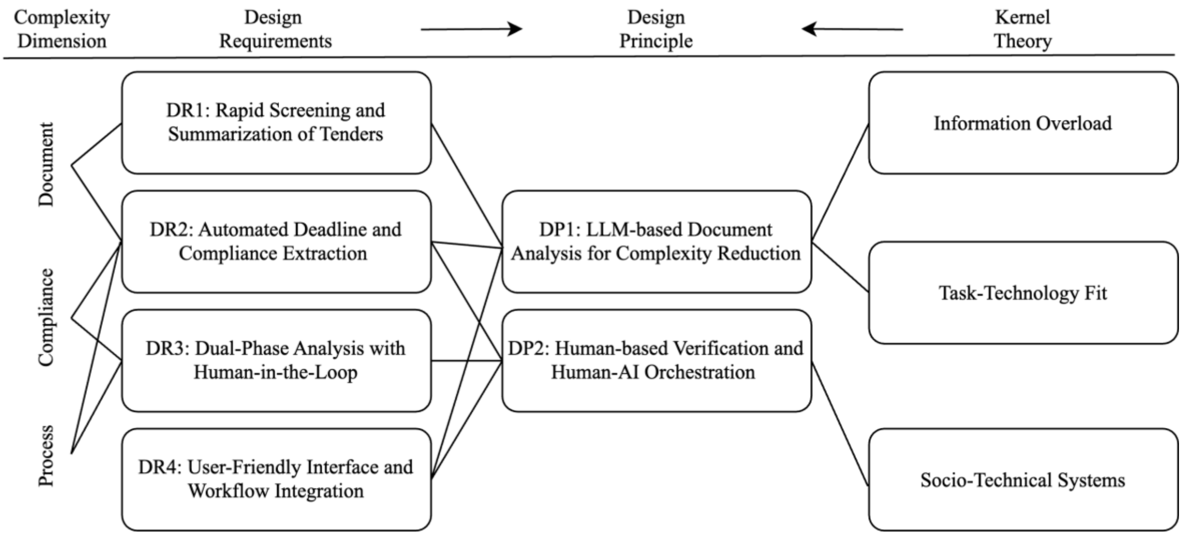

Figure 1. Mapping from design requirements and kernel theories to design principles.

4.2.3 Design and Development

We instantiated the DPs in an LLM-based tender analysis system. Users interact via a frontend, where they upload tender documents and access three views: Overview, Compliance, and Question & Answering. In the backend, documents are parsed, chunked, and stored in a vector database. LLM agents plan, implement, and evaluate three workflows that enable summarization, compliance, and requirement extraction, as well as question answering using a RAG approach. Within this architecture, three design features (DFs) realize the DPs.

DF1: Structured Tender Overview. DF1 supports rapid screening and reduced document complexity (DR1, DR4, DP1). After upload, the summarization agent retrieves representative segments and generates a concise overview, along with a structured header that includes key fields such as the procuring entity, scope, and submission deadline. Each retrieved piece of information is linked to the corresponding passage in the original document, allowing users to validate and verify the summary. This feature provides users with an immediate summary and overview of the tender, eliminating the need to read hundreds of pages. Interviewees consistently described the initial screening process as "finding a needle in a haystack" and reported experiencing information overload from long, heterogeneous documents. DF1 directly addresses this by externalizing document reading and information retrieval within the artifact, thereby reducing cognitive load for users.

DF2: Compliance Checklist and Timeline. DF2 addresses automated requirements and compliance extraction, as well as collaborative oversight (DR2, DR4, DP2). The compliance agent extracts formal obligations and requirements—deadlines, mandatory certifications, reference requirements, and exclusion criteria—and outputs a structured list. This is displayed to the user as a checklist and timeline. Each item retains a link to the originating passage in the document, allowing users to verify or correct the information retrieval. In practice, DF2 consolidates diffuse compliance information spread across annexes and documents into a single checklist. Our interviews revealed that missing a deadline or failing to meet a formal criterion results in immediate disqualification, and that requirements are often unclear and require further specification. DF2 lowers compliance risk and coordination effort by retrieving and displaying compliance requirements.

DF3: Source-Grounded Q&A. DF3 supports targeted information seeking that goes beyond DF1 and DF2 (DR1, DR3, DR4, DP1, DP2). Users can ask natural-language questions, and the Q&A agent retrieves relevant information and synthesizes it into an answer. Users can trace the information in the answer back to the original document and refine their questions iteratively. This feature enables users to examine complex documents conversationally, while still grounding and linking every answer to its original document, which the user can then verify. Interviewees emphasized the need for frequent ad hoc clarification of questions during tender analysis. DF3 answers these questions while enabling the user to verify the answer.

Together, these three DFs, as instantiated in our artifact, address document, compliance, and process complexity, shifting human work from information reading and extraction (read-and-extract) to orchestrating multiple LLM agents and verifying their outputs (orchestrate-and-verify).

4.3 Evaluation

Following Venable et al. (2016), we adopted a *Technical Risk & Efficacy* evaluation strategy for the proposed LLM-based tender analysis artifact. We conducted a quantitative evaluation of retrieval accuracy and time savings using real public tender documents, complemented by a qualitative think-aloud study with domain experts to examine perceived reductions in the complexity of tender analysis.

4.3.1 Quantitative Evaluation

To assess technical efficacy, we used a dataset of 34 real tender documents to measure retrieval accuracy and initial document screening time. Each tender was annotated with ground-truth labels for five structured text fields and two binary yes/no questions that capture core information needs in early screening: (1) content summary of the tender, (2) submission deadline for a bid proposal, (3) procuring entity, (4) place of execution, (5) type of service, and the binary classification questions (6) "required certifications mentioned?" and (7) "required references mentioned?". The dataset and labels were created in collaboration with the interviewed experts. The artifact was evaluated on all tender documents, and its outputs were compared with the manual annotations.

Information Retrieval Accuracy. The artifact demonstrated high accuracy in extracting the five text items, achieving an overall content-match accuracy of 88.82% (Figure 2, left). Accuracy is highest for the *procuring entity* and *place of execution* (both 94.10%) and for the *deadline* (91.20%), indicating that the artifact reliably captures key information that is typically expressed in relatively standardized parts of the tender. The *content summary* achieves an accuracy of 85.30%. The lower value compared to the other items is likely due to the open-ended nature of free-text summaries. The lowest accuracy is observed for the *type of service* at 79.40%, reflecting that this item is often described with heterogeneous, domain-specific wording and sometimes spread across multiple document pages.

Figure 2. Accuracy of retrieved text items (left) and binary classification task (right).

Binary Classification Accuracy. For the two binary classification tasks, the artifact achieved an average accuracy of 83.85% (Figure 2, right). Detection of *required certifications* achieved 91.20% accuracy. In contrast, the classification of *required references* reached 76.50% accuracy. This difference mainly arises when non-mandatory elements (such as “may provide references”) are interpreted as strict requirements, or when reference requirements are described indirectly, whereas certifications are typically stated more explicitly (e.g., *International Organization for Standardization* certifications).

Initial Screening Time. To assess time savings, we compared the initial screening time of domain experts with the processing time of the artifact. Manually, experts needed, on average, 3 minutes 42 seconds to screen one tender document and identify the seven retrieval items. The artifact processed a document in 1 minute 30 seconds on average, including uploading and indexing. This corresponds to an average time reduction of 2 minutes 12 seconds, i.e., approximately 59% faster than manual screening.

4.3.2 Qualitative Evaluation

To complement the quantitative evaluation and examine the perceived reduction in complexity in realistic use, we conducted a structured think-aloud study (Ericsson & Simon, 1993) with five domain experts. Each session lasted between 23 minutes and 47 minutes and consisted of (1) performing typical tender-analysis tasks with the artifact while verbalizing thoughts and (2) a short post-session interview guided by our three complexity dimensions (document, compliance, process) and the two DPs. During these sessions, the perceived reduction in complexity was assessed, supplemented by post-session interviews that were transcribed and deductively coded (Mayring, 2000). Several participants used their own current public tenders.

Document complexity. Participants consistently reported that the artifact reduced the effort required to obtain an initial overview of long, heterogeneous tender documents. They used the automated summary to understand scope, eligibility, and deadlines before turning to the full documents. The summary was described as a navigational aid, while linked passages enabled quick verification of information (one expert noted that they could “recognize the most important passages more quickly”). Participants also found the interface to be intuitive, supporting them in the rapid screening of documents without information overload, in line with the intention of LLM-based document analysis (DP1).

Compliance complexity. Regarding compliance complexity, participants primarily interacted with extracted deadlines, mandatory certifications, reference requirements, and exclusion criteria, which were presented as a structured checklist and timeline. This representation helped them identify potential omissions early, allocate follow-up tasks, and prepare clarifications where requirements were ambiguous or distributed across annexes. At the same time, participants emphasized that critical items would still be checked manually, and that the artifact should be used as a pre-screening aid rather than a fully autonomous compliance assessor. Minor inaccuracies in technical details were considered acceptable if the system provided a reliable orientation (one participant remarked that "small inaccuracies … are sufficient for an initial orientation"). The participants highlighted the value of the artifact, conducting a first automated screening of compliance-relevant content, while experts verify, correct, and interpret the outputs before making decisions. This thereby reduces perceived compliance complexity without removing human oversight (DP2).

Process complexity and work-mode shift. Participants positioned the artifact at the beginning of the tender analysis process as a screening step whose outputs (summary, compliance list, extracted items) can be handed over to technical, legal, and sales colleagues. They described initial familiarization with tender documents as one of the most time-consuming steps ("the step that often takes the longest is getting an initial overview") and that the artifact helped them clarify questions more quickly because relevant sections were highlighted, rather than having to be located manually. Four out of five participants explicitly named time savings as the primary benefit, particularly for the initial screening phase and the bid/no-bid decision. They emphasized that access to technical, legal, and administrative details enabled them to quickly locate relevant information and establish a common starting point for departments to collaborate. Taken together, these observations support a shift from reading and manually extracting information to orchestrating the artifact for screening and then verifying or refining its outputs (DP1, DP2).

In summary, the qualitative evaluation supports the notion that the instantiated DPs in our artifact reduces complexity. (1) Document complexity is reduced by transforming long, heterogeneous documents into concise overviews that facilitate rapid screening. (2) Compliance complexity is reduced by structuring requirements into checklists and timelines, enabling experts to spot omissions early and focus their verification on critical items. (3) Process complexity is reduced by providing information that serves as a starting point for sales, legal, and technical roles, thereby shortening initial familiarization and accelerating bid/no-bid decisions. Overall, the findings indicate that participants aligned with the intended orchestrate-and-verify work mode of DP1 and DP2, in which LLM-based analysis provides a structured starting point and human experts coordinate and validate the results.

5 Discussion

This study examined how LLM-based IS can be designed to manage and reduce complexity in document-intensive, compliance-critical knowledge work, using bidder-side public tender analysis in an SME. Building on interviews, we identified three interconnected dimensions of complexity: document complexity (volume, heterogeneity, and fragmentation), compliance complexity (strict, sometimes ambiguous formal requirements and disqualification risks), and process complexity (cross-departmental coordination under time constraints). We derived four DRs and two DPs for complexity-reducing human–AI systems. We instantiated these principles in an LLM-based tender analysis artifact that automates summarization, compliance extraction, and source-grounded Q&A to support tender document screening and early bid/no-bid decisions. In a quantitative evaluation on real tender documents, the artifact reduced initial screening time by 59% while maintaining high retrieval accuracy (88.82% for text items and 83.85% for binary checks). A think-aloud study showed that experts perceived reductions in complexity and shifted from manual reading and extraction toward an orchestrate-and-verify work mode. In the following sections, we discuss how these findings contribute to IS research, human–AI collaboration, LLM-based systems, and discuss practical implications.

5.1 Theoretical Implications

Complexity as a Multi-Dimensional Design Challenge. Our findings affect IS research on complexity management and reduction in knowledge work by showing how complexity in document-intensive, compliance-critical work can be addressed. Prior work has emphasized information overload and coordination demands as key sources of complexity in organizational settings (Eppler & Mengis, 2008; Okhuysen & Bechky, 2009; Voglhofer & Rinderle-Ma, 2020; Wunderlich & Fischer, 2022). We extend this line of work by empirically demonstrating that complexity manifests in three distinct yet interrelated dimensions—document, compliance, and process complexity—and by mapping each dimension to concrete challenges faced by knowledge workers, like non-standard document formats, compliance risks, and cross-departmental coordination. Rather than treating complexity as a fixed environmental condition, our study contributes design knowledge on how complexity can be reduced and shifted through the design of human–AI systems. We emphasize that complexity is not eliminated; instead, it shifts from reading and manually extracting information to orchestrating LLM-based IS and verifying their outputs, which, from a human-AI collaboration perspective, contain their own complexities and challenges. This complements recent IS research that examines how digitalization reshapes task complexity in knowledge work, and that LLM-based systems can both reduce and shift complexity, rather than assuming that automation automatically simplifies work (Wunderlich & Fischer, 2022). The transferability of the three complexity dimensions and the associated DPs is bounded by knowledge work contexts that combine heterogeneous documents, consequential formal requirements, and coordination across multiple expert roles. Future IS research should develop methods that measure and theorize such complexity shifts, investigate them in different contexts, and create design knowledge to support systems for humans when complexity shifts occur.

Human-AI Orchestration as a New Mode of Knowledge Work. A second implication concerns the mode of human–AI collaboration that emerges when LLMs are used for complexity management in document-intensive, compliance-critical work. Our evaluation indicates that GenAI for knowledge workers does not simply eliminate complexity or replace humans entirely. Instead, a new human-AI work mode emerges, which we term the orchestrate-and-verify work mode. In this mode, knowledge workers (1) interact with and orchestrate one or multiple LLM-based systems for different document and information analysis tasks, such as retrieving key information into structured overviews, checklists, and timelines. Knowledge workers then (2) verify and correct the outputs using source links and expert knowledge. This dual-phase work mode is encoded in our design knowledge and reflected in practice when participants use the artifact to obtain an initial overview, ask targeted questions, and then selectively drill down into the underlying passages to verify information they deem critical. This resonates with emerging IS and future-of-work perspectives that see GenAI as reallocating cognitive labour from routine tasks toward problem framing, critical evaluation, and oversight (Feuerriegel et al., 2024; Söllner et al., 2025). Additionally, this aligns with a human-centred vision for GenAI systems that emphasises high levels of both human control and computer automation to enhance performance while maintaining responsibility (Shneiderman, 2020), as well as with work on human–AI interaction that recommends designs enabling users to understand, control, and correct AI behaviour (Amershi et al., 2019). LLMs further make collaboration between humans and AI increasingly interactive, as they allow users to iteratively query, refine, and redirect GenAI systems through natural language rather than only receiving fixed predictions or recommendations (Rajashekar et al., 2024). Effective collaboration and orchestration depend on humans continuously updating their understanding of AI capabilities, limitations, and complementarity with human expertise (Holstein & Satzger, 2025). Existing literature, therefore, already shows that GenAI can augment knowledge work, reallocate cognitive effort toward oversight, and require human control and correction. What remains less clear, however, is how these dynamics translate into a concrete mode of collaboration. Additionally, verification becomes crucial, because LLM outputs and self-explanations are not reliably faithful and outputs may still contain omissions or errors (Zhang et al., 2025). Humans cannot rely on LLM-generated outputs alone but must verify and, where necessary, correct them through source links and expert judgment. Thus, our study contributes not by showing that GenAI augments work or requires oversight in general, but by explaining how the combination of interactional capability and error risk reorganizes document-intensive, compliance-critical knowledge work into an orchestrate-and-verify mode. This is likely to become even more important as recent research increasingly turns to multi-agent AI and agentic systems, in which users must coordinate and orchestrate not only single model outputs, but interacting agents (Allmendinger et al., 2026).

From a critical perspective, this shift raises at least three concerns. (1) Deskilling and shallow understanding: When GenAI systems consistently provide pre-digested overviews, knowledge workers may read fewer primary documents in depth. Over time, this could erode tacit domain knowledge and reduce the ability to spot subtle issues that the model misses. Our study cannot yet determine whether such deskilling occurs, but we highlight the risk that short-term efficiency gains may come at the expense of long-term expertise. (2) Diffuse accountability: If knowledge workers rely on LLM-generated overviews, accountability for decisions may become diffuse as LLMs may miss key information or present false information. Human-centred AI frameworks insist on clear lines of responsibility (Shneiderman, 2020); yet many current GenAI systems, including ours, rely on informal norms (e.g., "always verify") rather than formal governance. (3) New forms of cognitive pressure: While GenAI reduces some forms of complexity, it introduces new ones: monitoring model behaviour, reconciling conflicting outputs, and deciding when to trust or override the system. Users may experience a subtler, yet still significant, kind of mental strain and complexity: spending less time reading documents and more time second-guessing an opaque model. For IS, this points to a broader research agenda: we need theories and design knowledge for orchestrate-and-verify work that address expertise development, calibration of trust, and responsibility.

LLM-Based Systems as a Socio-Technical Design Pattern. Our findings demonstrate that LLMs function as complexity redistributors rather than complexity eliminators, fundamentally reconfiguring the division of labor between humans and machines in knowledge work. In tender analysis, we observe that complexity does not disappear but shifts its form: workers no longer spend extensive time reading heterogeneous PDFs, but instead orchestrate LLM agents and validate their extractions; compliance complexity transforms from manual deadline tracking across dispersed documents to reviewing machine-generated checklists; process complexity shifts from fragmented cross-departmental information gathering to coordinating pre-analyzed summaries. This pattern only emerged when multiple socio-technical elements aligned: LLM retrieval capabilities, source-linking mechanisms that enable verification, domain expertise to identify errors, organizational workflows that separate screening from detailed analysis, and institutional tolerance for probabilistic outputs. These findings have important implications for how we operationalize LLM-based systems in knowledge work. Rather than viewing them as standalone technical components that automate tasks, they should be understood as socio-technical design patterns for complexity management where effectiveness depends on deliberate orchestration across technology, work practices, expertise, and institutional constraints. This perspective reveals systemic risks invisible to purely technical evaluations: while retrieval accuracy and time reduction demonstrate efficiency gains, these metrics obscure potential long-term deskilling as workers read fewer primary documents, diffuse accountability when the system misses critical requirements, and uneven adoption barriers where digitally mature SMEs benefit while others fall further behind in public procurement participation. Unlike prior LLM applications that automate isolated subtasks, our design pattern demonstrates that the nature of work fundamentally changes from information extraction to output verification—a shift demanding new skills, governance structures, and risk frameworks that extend beyond technical performance alone.

5.2 Practical Implications

Strategic Repositioning of GenAI from Automation to Complexity Management. For practitioners in public procurement and other document-intensive, compliance-critical domains, our findings suggest repositioning GenAI from a tool for full automation towards a system for complexity management. Rather than expecting LLM-based systems to make decisions autonomously, organizations should evaluate and adopt them based on their ability to reduce and redistribute document, compliance, and process complexity in early stages—for instance, by providing structured overviews, obligation lists, and shared views that support cross-departmental coordination. Positioning systems explicitly as screening and orientation aids helps align stakeholder expectations, mitigate fears of replacement, and focus governance on where human expertise must remain central (e.g., final bid decisions, interpreting ambiguous requirements). This implies framing GenAI initiatives within organizations as complexity-reducing and complexity-redistributing projects that complement, rather than replace, experts.

Rethinking Workforce Skills, Training, and AI Literacy. The orchestrate-and-verify work mode has direct implications for workforce development. Instead of primarily training experts to read and interpret entire documents manually, organizations need to cultivate skills for orchestrating and critically supervising LLM-based IS, including formulating suitable requests, understanding how systems retrieve and generate content, interpreting confidence signals and source links, and systematically verifying model outputs. Our findings suggest that such training should explicitly address complexity redistribution—helping experts recognize which parts of the document and compliance complexity are shifted to the system and which new responsibilities arise on the human side. This includes understanding model limitations, knowing when additional human investigation is necessary, and documenting how GenAI outputs are used in decision processes. Organizations that invest early in these competencies are likely to be better positioned to leverage GenAI while avoiding overreliance, compliance risks, and long-term deskilling. This implies that GenAI does not eliminate the need for expert knowledge work; instead, it reconfigures it, increasing the importance of oversight, critical thinking, and AI literacy (Budhwar et al., 2023), beyond merely understanding how GenAI works, towards a holistic and human-centered understanding of the impact of GenAI in knowledge work.

5.3 Limitations and Future Work

This study has several limitations that point to future research. First, our work is conducted in collaboration with a single SME in the data centre sector and in bidder-side tender analysis within a specific regulatory regime; complexity reduction may partly reflect novelty effects and enthusiasm for GenAI, and our three complexity dimensions and DPs may not transfer unchanged to other industries or types of knowledge work. We therefore do not claim direct generalizability across all SMEs or sectors. Rather, transfer is most plausible in contexts that share similar knowledge work characteristics like dealing with large heterogeneous document collections, as in law, healthcare, or the construction industry (Diener et al., 2025). Future research should therefore extend our findings to organizations of different sizes, sectors, and countries, as well as to non-procurement settings, while also capturing organizational-level dynamics and practical adoption barriers, like initial setup costs, for resource-constrained SMEs. Second, our evaluation captures short-term use with a limited number of documents and experts, providing only an initial view of how complexity is redistributed and how the work mode is adopted over time. Long-term, experimental studies are needed to research how knowledge work practices evolve and how initial complexity reductions are sustained or offset by new orchestration and verification demands of LLM-based IS. Third, to help humans understand the new complexities of the orchestrate-and-verify work mode, future research should investigate the types of errors made by LLM-based IS and also compare this approach with more traditional technologies.

6 Conclusion

This paper examined how LLM-based systems can be designed to reduce complexity in document-intensive, compliance-critical knowledge work, using bidder-side public tender analysis in a German SME as our focal context. Applying a DSR approach, we identified and operationalized complexity for knowledge workers along document, compliance, and process dimensions, derived four DRs, and formulated two DPs for LLM-based IS. We instantiated these principles in an LLM-based tender analysis artifact and evaluated it with real public tenders and domain experts, showing that the artifact reduces initial screening time while maintaining high retrieval accuracy and that experts experience reduced complexity and a shift in their work practices. Overall, our contributions are threefold. First, we contribute to IS research on complexity management by operationalizing complexity in document-intensive knowledge work as the interplay of document, compliance, and process complexity. Second, we provide prescriptive design knowledge in the form of four DRs and two DPs for LLM-based IS and demonstrate their instantiation and efficacy in an artifact that combines structured overviews, compliance checklists, timelines, and question answering. Third, we extend research on LLM-based IS by articulating human-based verification and human-AI orchestration as a new mode of knowledge work, in which LLM-based systems redistribute rather than eliminate complexity.

7 References

Allmendinger, S., Bonenberger, L., Endres, K., Fetzer, D., Gimpel, H., & Kühl, N. (2026). Multi-agent AI. *Electronic Markets*, *36*(1), 18.

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., & others. (2019). Guidelines for human-AI interaction. *Proceedings of the 2019 Chi Conference on Human Factors in Computing Systems*, 1–13.

Benbya, H., Strich, F., & Tamm, T. (2024). Navigating Generative Artificial Intelligence Promises and Perils for Knowledge and Creative Work. *Journal of the Association for Information Systems*, *25*, 23–36.

Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., Boselie, P., Lee Cooke, F., Decker, S., DeNisi, A., Dey, P. K., Guest, D., Knoblich, A. J., Malik, A., Paauwe, J., Papagiannidis, S., Patel, C., Pereira, V., Ren, S., … Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal*, *33*(3), 606–659.

Davis, F. D., & Granić, A. (2024). *The Technology Acceptance Model: 30 Years of TAM*. Springer International Publishing.

Diener, M., Schäfer, S., Spitzer, P., & Vössing, M. (2025). Constructing Insights: Leveraging Large Language Models for Information Retrieval in Unstructured Document Collections with ConSight. *20th International Conference on Wirtschaftsinformatik (WI2025)*.

Eppler, M. J., & Mengis, J. (2008). The Concept of Information Overload—A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines (2004). In M. Meckel & B. F. Schmid (Eds.), *Kommunikationsmanagement im Wandel: Beiträge aus 10 Jahren =mcminstitute* (pp. 271–305). Gabler.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data.* The MIT Press.

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization, 28*(1), 62-70.

Fazekas, M., & Blum, J. R. (2021). *Improving Public Procurement Outcomes: Review of Tools and the State of the Evidence Base* (Policy Research Working Paper No. 9690). World Bank. http://hdl.handle.net/10986/35727

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

Flynn, A. (2017). Re-thinking SME disadvantage in public procurement. *Journal of Small Business and Enterprise Development*, *24*(4), 991–1008.

Flynn, A. (2025). Research on SME involvement in public procurement: A review, critique and conceptual framework. *Journal of Purchasing and Supply Management*, 101052.

Flynn, A., & Davis, P. (2017). Explaining sme participation and success in public procurement using a capability-based model of tendering. *Journal of Public Procurement*, *17*(3), 337–372.

Flynn, A., McKevitt, D., & Davis, P. (2015). The impact of size on small and medium-sized enterprise public sector tendering. *International Small Business Journal*, *33*(4), 443–461.

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). *Will Humans-in-The-Loop Become Borgs? Merits and Pitfalls of Working with AI* (SSRN Scholarly Paper No. 3879937). Social Science Research Network. https://papers.ssrn.com/abstract=3879937

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, *16*(1), 15–31.

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, *19*(2), 213–236.

Gregor, S., Kruse, L. C., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems*, *21*, 1622–1652.

Harder Fischer, L., Nicolajsen, H. W., Marttila, S.-M., & Sandbukt, S. (2024). Crafting Meaningful Generative AI-Enabled Knowledge Work. *ECIS 2024 Proceedings*. https://aisel.aisnet.org/ecis2024/track01_peoplefirst/track01_peoplefirst/10

Hassani, S. (2024). Enhancing Legal Compliance and Regulation Analysis with Large Language Models. *2024 IEEE 32nd International Requirements Engineering Conference (RE)*, 507–511.

Holstein, J., & Satzger, G. (2025). Development of Mental Models in Human-AI Collaboration: A Conceptual Framework. *ICIS 2025 Proceedings*. https://aisel.aisnet.org/icis2025/hti/hti/2

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, *66*(1), 87–99.

Kaiser, R. (2021). *Qualitative Experteninterviews: Konzeptionelle Grundlagen und praktische Durchführung*. Springer Fachmedien.

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, *72*(12), 2954–2965.

Kirchner, K., Bolisani, E., Kassaneh, T. C., Scarso, E., & Taraghi, N. (2025). Generative AI Meets Knowledge Management: Insights From Software Development Practices. *Knowledge and Process Management*, *32*(4), 223–235.

Klesel, M., & Wittmann, H. F. (2025). Retrieval-Augmented Generation (RAG). *Business & Information Systems Engineering*, *67*(4), 551–561.

Krakowski, S. (2025). Human-AI agency in the age of generative AI. *Information and Organization*, *35*(1), 100560.

Kudina, O., & van de Poel, I. (2024). A sociotechnical system perspective on AI. *Minds and Machines*, *34*(3), 21.

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: Anatomy of a research project. *European Journal of Information Systems*, *17*(5), 489–504.

Loader, K. (2015). SME suppliers and the challenge of public procurement: Evidence revealed by a UK government online feedback facility. *Journal of Purchasing and Supply Management*, *21*(2), 103–112.

Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-Line Journal], Http://Qualitative-Research.Net/Fqs/Fqs-e/2-00inhalt-e.Htm*, *1*.

McCue, C., & Roman, A. V. (2012). E-Procurement: Myth or Reality. *Journal of Public Procurement*, *12*(2), 221–248.

Morrison, K., Iqbal, S. T., & Horvitz, E. (2024). AI-Powered Reminders for Collaborative Tasks: Experiences and Futures. *Proc. ACM Hum.-Comput. Interact.*, *8*(CSCW1), 210:1-210:20.

Mumford, E. (2006). The story of socio-technical design: Reflections on its successes, failures and potential. *Information Systems Journal*, *16*(4), 317–342.

Okhuysen, G. A., & Bechky, B. A. (2009). Coordination in Organizations: An Integrative Perspective. *The Academy of Management Annals*, *3*(1), 463–502.

Oussaleh Taoufik, A., & Azmani, A. (2023). *Smart Sourcing Framework for Public Procurement Announcements Using Machine Learning Models* (pp. 921–932).

Rajashekar, N. C., Shin, Y. E., Pu, Y., Chung, S., You, K., Giuffre, M., Chan, C. E., Saarinen, T., Hsiao, A., Sekhon, J., Wong, A. H., Evans, L. V., Kizilcec, R. F., Laine, L., Mccall, T., & Shung, D. (2024). Human-Algorithmic Interaction Using a Large Language Model-Augmented Artificial Intelligence Clinical Decision Support System. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI '24*, 1–20.

Roetzel, P. G. (2019). Information overload in the information age: A review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research*, *12*(2), 479–522.

Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Three Fresh Ideas. *AIS Transactions on Human-Computer Interaction*, *12*(3), 109–124.

Söllner, M., Arnold, T., Benlian, A., Bretschneider, U., Knight, C., Ohly, S., Rudkowsi, L., Schreiber, G., & Wendt, D. (2025). ChatGPT and Beyond: Exploring the Responsible Use of Generative AI in the Workplace. *Business & Information Systems Engineering*, *67*(2), 289–303.

Statistisches Bundesamt (Destatis). (2025). *Public contracts and concessions awarded, volume of public contracts and concessions: Germany, quarters, contracting authority level, threshold, type of contract*. https://www-genesis.destatis.de/datenbank/online/url/4b792a35

Storey, V. C. (2025). Knowledge Management in a World of Generative AI: Impact and Implications. *ACM Trans. Manage. Inf. Syst.*, *16*(3), 26:1-26:14. https://doi.org/10.1145/3719209

Trist, E. L. (1981). *The evolution of socio-technical systems* (Vol. 2). Ontario Quality of Working Life Centre Toronto.

Vaidya, K., Sajeev, A. S. M., & Callender, G. (2006). Critical factors that influence e-procurement implementation success in the public sector. *Journal of Public Procurement*, *6*(1/2), 70–99.

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, *25*(1), 77–89.

Voglhofer, T., & Rinderle-Ma, S. (2020). Collection and Elicitation of Business Process Compliance Patterns with Focus on Data Aspects. *Business & Information Systems Engineering*, *62*(4), 361–377.

Vuori, V., Helander, N., & Okkonen, J. (2019). Digitalization in knowledge work: The dream of enhanced performance. *Cognition, Technology & Work*, *21*(2), 237–252.

Wunderlich, N., & Fischer, L. (2022). How Digital Business Strategy Shapes Task Complexity in Datafication of Knowledge Work. *Hawaii International Conference on System Sciences 2022 (HICSS-55)*. https://aisel.aisnet.org/hicss-55/os/design/4

Zhang, Y., Li, Y., Cui, L., Cai, D., Liu, L., Fu, T., Huang, X., Zhao, E., Zhang, Y., Chen, Y., Wang, L., Luu, A. T., Bi, W., Shi, F., & Shi, S. (2025). 🧜Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models. *Computational Linguistics*, *51*(4), 1373–1418.