

Schryen, Guido; Marrone, Mauricio; Yang, Jiaqi

Article — Published Version

Exploring the scope of generative AI in literature review development

Electronic Markets

Provided in Cooperation with:

Springer Nature

Suggested Citation: Schryen, Guido; Marrone, Mauricio; Yang, Jiaqi (2025) : Exploring the scope of generative AI in literature review development, Electronic Markets, ISSN 1422-8890, Springer, Berlin, Heidelberg, Vol. 35, Iss. 1, <https://doi.org/10.1007/s12525-025-00754-2>

This Version is available at:

<https://hdl.handle.net/10419/319041>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Exploring the scope of generative AI in literature review development

Guido Schryen¹ · Mauricio Marrone² · Jiaqi Yang^{2,3}

Received: 4 April 2024 / Accepted: 7 January 2025
© The Author(s) 2025

Abstract

Artificial intelligence (AI) has the potential to transform the way research is conducted, particularly through generative AI (GenAI) tools which can enhance written communication and foster innovation via knowledge development. This study focuses on the latter, examining the role of GenAI in specific knowledge development activities within literature reviews. Through an epistemological lens, we distinguish six key knowledge development activities: research synthesis, evidence aggregation, critique, theory building, research gap identification, and research agenda development. Our analysis demonstrates both the capabilities and limitations of GenAI in supporting these activities, highlighting how GenAI can assist in synthesizing previous work, discovering and integrating concepts, and advancing various knowledge domains. We emphasize a human-centered, synergistic approach where GenAI complements researchers' efforts, rather than replacing them. Additionally, our activity-centric analysis provides insights into how different types of literature reviews can effectively benefit from GenAI support, thereby contributing to a broader understanding of AI integration in information systems research.

Keywords Generative AI · Literature reviews · Knowledge development · Innovation goal

JEL Classification M00

Introduction

Generative artificial intelligence (GenAI) is a highly potent sub-category of artificial intelligence (AI) that has gained considerable prominence, largely due to exemplars such as ChatGPT. GenAI operates by leveraging deep learning models to generate human-like content, such as images and

words, in response to complex and diverse linguistic inputs, instructions, or inquiries (Lim et al., 2023).

GenAI tools, such as ChatGPT and Google Gemini (formerly known as Bard), have the potential to enhance scholarly work. For example, Visconti (2021) creates a machine-generated literature review for climate, planetary, and evolutionary sciences. The capabilities of AI tools are rapidly evolving, often surpassing our predictions. In terms of academic research, they could achieve primary goals: the improvement of writing (communication goal) and the generation of new ideas (innovation goal) (Dwivedi et al., 2023). Focusing on the communication goal, GenAI tools like ChatGPT can aid in proofreading, editing, and refining the writing of the papers. They complement existing writing tools, such as Grammarly and Spellcheck, which are particularly beneficial for non-native English-speaking researchers. GenAI can improve language quality and clarity, ensuring that complex ideas are communicated effectively. Many scholars provide preliminary feedback regarding the use of GenAI tools for scientific communication, often making recommendations and expounding best practices. For example, readers can refer to the works of Buriak et al. (2023), van Dis et al. (2023), and Schlagwein and Willcocks (2023).

Responsible Editor: Rainer Schmidt

✉ Guido Schryen
guido.schryen@upb.de
Mauricio Marrone
mauricio.marrone@mq.edu.au
Jiaqi Yang
jiaqi.yang@hdr.mq.edu.au

¹ Department of Management Information Systems, Paderborn University, Warburger Str. 100, 52062 Aachen, Germany

² Present Address: Macquarie Business School, Macquarie University, 4 Eastern Road, Sydney 2113, Australia

³ School of Information Systems and Technology Management, UNSW Sydney, E15 College Rd, Kensington NSW 2033 Sydney, Australia

Although employing GenAI to enhance research communication is relatively straightforward, leveraging it to achieve research innovation is complex and has generated considerable debate. The innovation goals highlight GenAI's role in exploring and generating ideas, integrating multidisciplinary perspectives, solving research problems creatively, and proposing new theoretical insights (Dwivedi et al., 2023). Unlike communication goals, which highlight GenAI's role in scholarly presentation, innovation goals emphasize GenAI's role in various knowledge activities. From the innovation goals, we derive our understanding that knowledge development is one important type of innovation, as it involves the continuous creation, refinement, and integration of existing and new concepts to advance the knowledge domain. On one hand, GenAI holds the potential to deeply engage in the knowledge development process by contributing to selecting theoretical products, identifying focal ideas, and establishing theory-building apparatus (Jarvenpaa & Klein, 2024). It can be used to uncover insights that are not immediately obvious to researchers, serving as stimuli for novel ideas and encouraging the exploration of new knowledge (Benbya et al., 2024). On the other hand, numerous challenges related to GenAI tools in knowledge development, including hallucination, interpretability, and institutionalization biases, are well noted (Susarla et al., 2023). These tools, based on generative textual engines, are designed to rely on words and phrases from their training data, rather than on logic, semantic, or epistemic models. As a result, they have been described as “stochastic parrots” that build sentences from data traces (Bender et al., 2021). Moreover, GenAI's reliance on past data and inability to grasp subjective experiences or context may restrict its ability to develop new ideas; it frequently perpetuates outdated practices, which can lead to misinformation and stifle innovation in knowledge development (Benbya et al., 2024). Therefore, pursuing the innovation goal with GenAI requires us to delve deeper into its suitability for supporting specific knowledge activities.

We acknowledge that knowledge activities encompass a variety of research paradigms, methods, and genres—such as knowledge creation and generation (Alavi & Leidner, 2001), knowledge capture and discovery (Paul, 2006), knowledge integration and synthesis (Majchrzak et al., 2013), and knowledge refinement and evolution (Ramakrishnan et al., 2023). Given the breadth and complexity of these knowledge activities, a comprehensive exploration of GenAI's role across all domains would be beyond the scope of a single paper. Instead, this work focuses on the dedicated knowledge development activities within literature reviews as a typical element of (almost) all research endeavors and publications, including “regular” research papers and standalone literature reviews. The particular importance of literature reviews in the context of knowledge development lies in the understanding that every literature review generates some knowledge through its mandatory activity of synthesizing

previous work (Schryen et al., 2020). This synthesizing activity involves not only the discovery of knowledge through a typically structured literature search and evaluation process but also the process of describing concepts and using them to integrate, relate, contrast, and organize the discovered knowledge in a concept-centric manner (Schryen, 2015; Webster & Watson, 2002). This process requires creativity and human understanding of the discovered body of knowledge, resulting in a new knowledge contribution of its own; i.e., even those literature reviews that merely synthesize prior knowledge also develop new knowledge through the formulation of concepts and their use to present discovered knowledge.

Overall, the question of how to support knowledge development in literature reviews with GenAI is relevant to most researchers. From this focus, we derive our central question, which we seek to answer “How can GenAI tools be used effectively to support specific knowledge development activities in information systems (IS) literature reviews?” Our focus is to explore how the use of GenAI can provide methodological support and foster knowledge development in literature reviews (as standalone reviews or parts of other research works) in a human-AI collaboration.

To address our research question, we take an epistemological perspective on literature reviews and draw on a widely adopted set of knowledge development activities, including synthesizing (including discovering) prior research, criticizing prior research, aggregating evidence, theory building, identifying research gaps, and developing a research agenda. Our goal is to evaluate the suitability of GenAI tools for conducting these activities, and from this analysis, we then offer recommendations of effectively using GenAI tools. However, it should be noted that the epistemological perspective on GenAI is different from an analysis of how the increasingly sophisticated technical capabilities of GenAI tools can be used to process (e.g., summarize, extract, compare, consolidate, modify) text documents, spreadsheets, images, videos, audio, etc. While such activities refer to a predefined set of input and can mostly be performed without human intervention, knowledge development activities refer to the vast body of all training data of the GenAI LLM and require some form of human-AI collaboration. It should be further clarified that our study does not aim to examine how different GenAI tools respond to different queries; our study is also neither confirmatory nor exploratory. Instead, we draw on the epistemological nature of knowledge development activities and the principles of GenAI tools, and we utilize examples of GenAI queries to showcase our recommendations and implications. Through this, we aim to demonstrate the potential of GenAI tools in human-AI collaboration for compiling IS LRs and to suggest strategies for enhancing the efficiency of the research process and improving the quality of the research results.

Background

Literature reviews through an epistemological lens

The genre of literature review has attracted much interest in many scientific fields for decades, resulting in various classifications of LRs. In the IS discipline, researchers have developed typologies that classify LRs according to their research goals and methods. For example, Rowe (2014) distinguishes four goals (describing, explaining, understanding, and theory testing). In line with these goals, Paré et al. (2015) distinguish nine types of LRs that synthesize prior knowledge, aggregate or integrate data, construct explanations, or assess extant literature critically. An epistemological perspective to distinguish LRs has been proposed by Schryen et al. (2020). Table 1 shows which knowledge development¹ activities can be identified in LRs and how they align with the above typologies.

As can be seen from Table 1, six knowledge development activities are distinguished: synthesizing research (SYN), aggregating evidence (AE), criticizing (CRI), theory building (TB), identifying research gaps (RG), and developing a research agenda (RA). While the activities SYN, AE, and CRI focus on past research and can be considered backward-oriented, the activities RG, RA, and TB point to future research and are thus forward-oriented. As we draw on these activities to discuss whether and how GenAI can be used to support knowledge development in LRs, Table 2 briefly explains the essence of these activities.

Principles of GenAI

GenAI broadly refers to a class of AI models that produces seemingly new and meaningful content in the form of text, images, or other media. These models function by learning patterns from their extensive training datasets and generate content based on those patterns (Sursala et al., 2023). Notable GenAI tools include Gemini and ChatGPT as Generative Language Models, Dall-E 3, Stable Diffusion and Sora as Generative Image and Video Models, and Perplexity as an AI-enhanced Search Engine. The release of tools like GPT-4o and Dall-E 3 allows the generation of human-like output in text and visual formats with great sophistication. At the heart of these advancements are deep neural networks and transformational architectures as well as the availability of vast amounts of training data, which enable these models to predict and generate content in ways that closely mimic

human language, similar to an advanced form of autocompletion technology (Feuerriegel et al., 2024), visual artifacts developed by humans, including real-world pictures, scientific illustrations, artistic paintings, and other media.

Despite the impressive capabilities of GenAI, significant limitations exist based on using, training, and applying LLMs and the resulting way in which information is processed and output is generated. While GenAI can provide information, summarize extensive material, and generate coherent text, it does not “understand” the material in the human sense. For instance, AI can summarize research findings or explain concepts as described in the prompts, but it does not inherently grasp the underlying principles, context, or importance of these concepts beyond their textual representation. In light of this, scholars debate whether GenAI tools genuinely “understand” their outputs (Mitchell et al., 2023). AI systems can recognize patterns and reproduce data based on probabilities, but they lack the ability to contextualize knowledge within broader philosophical or theoretical frameworks unless explicitly outlined in the data they process. GenAI does not possess the human-like ability to creatively theorize or speculate in a way that reflects deep understanding and innovative thought. Also, of particular importance is the inability of GenAI to include or simulate “human intuition,” which goes beyond accessing and reproducing data. It lacks the ability to question deeply, think critically, and engage with material in a way that challenges and extends existing knowledge boundaries. Human intuition and experiential learning play crucial roles in these processes, allowing scholars to discern/disentangle subtleties and implications that might not be evident through AI. As a consequence of the described GenAI-inherent characteristics, we still need researchers “to know” and a collaborative human AI working model.

One resulting major issue of the aforementioned GenAI principles is the propensity for these models to produce incorrect or misleading results (often called “AI hallucinations”), where generated content appears plausible but is factually incorrect or nonsensical (Hicks et al., 2024). This problem is rooted in the probabilistic nature of these models, which generate the most likely response to a prompt, rather than verifying its truthfulness (Feuerriegel et al., 2024). Additionally, GenAI models frequently exhibit biases reflective of the human-generated data they are trained on, perpetuating stereotypes and prejudices present in the training data (Bail, 2024). Copyright violations also pose a significant limitation, as GenAI models can produce outputs that resemble existing works without permission or attribution to the original creators (Feuerriegel et al., 2024). Addressing these limitations requires ongoing research to improve model transparency, bias mitigation, and the development of ethical guidelines for AI deployment.

¹ While Schryen et al. (2020) use the term “knowledge-building activities,” we prefer to use the term “knowledge development activities” to remain terminologically consistent with the common term “knowledge development.”

Table 1 Derivation of knowledge development activities of literature reviews from methodological papers (Schryen et al., 2020, p. 137)

Goals ^a	Review types ^b	Knowledge development activity ^c	
		Backward-oriented	Forward-oriented
Describing	Narrative Review Levy and Ellis (2006), Hart (2009)	<ul style="list-style-type: none"> • Narrative summary of prior findings on a topic (SYN) 	<ul style="list-style-type: none"> • Identification of research gaps (RG) • Development of an agenda for research and practice (RA)
	Descriptive Review King and He (2005)	<ul style="list-style-type: none"> • Quantitative and narrative summary of what we know about a topic (SYN) • Identification of trends over time (SYN) 	<ul style="list-style-type: none"> • Development of recommendations to influence the development of a topic, domain, or method (RA)
	Scoping Review Arksey and O'Malley (2005), Levac et al. (2010)	<ul style="list-style-type: none"> • Narrative summary of the size and nature of extant literature (SYN) 	<ul style="list-style-type: none"> • Identification of research gaps (RG) • Development of a research agenda with potential implications for research and practice (RA)
Understanding	Critical Review Rowe (2014), Alvesson and Sandberg (2011)	<ul style="list-style-type: none"> • Summarize past knowledge on a domain of interest (SYN) • Critical account of extant literature, revealing weaknesses, or inconsistencies (CRI) 	<ul style="list-style-type: none"> • Providing a focus or a new direction to studies (RA)
	Theoretical Review Rivard (2014), Rowe (2014), Torraco (2005), Walker and Avant (2011), Webster and Watson (2002)	<ul style="list-style-type: none"> • Synthesis of prior literature (SYN) 	<ul style="list-style-type: none"> • Theory derivation: Development of a theory from the explanations in another field (TB) • Theory synthesis: Development of a theory from pulling together prior evidence about a phenomenon (TB) • Theory analysis: Examination of a theory and identification of the need for additional refinement (TB) • Developing a research agenda (RA)
Theory testing	Realist Review Pawson et al. (2005)	<ul style="list-style-type: none"> • Synthesis of evidence and dissemination of findings (SYN) 	<ul style="list-style-type: none"> • Development of a theory aimed at explaining what it is about an intervention that works, for whom, in what circumstances and why (TB) • Exploration of moderators can provide forward-looking ideas for future research (RA)
	Meta-analysis King and He (2005), Rosenthal and DiMatteo (2001), Card (2011)	<ul style="list-style-type: none"> • Integration of knowledge gained in empirical studies (SYN) • Statistical aggregation of empirical findings (AE) 	
	Qualitative Systematic Review Gough et al. (2012), Petticrew and Roberts (2008)	<ul style="list-style-type: none"> • Synthesis of evidence (SYN) • Narrative aggregation of possibly heterogeneous empirical findings (AE) 	<ul style="list-style-type: none"> • Development of implications for policy, practice, and further research (RG)
	Umbrella Review Thomson et al. (2010)	<ul style="list-style-type: none"> • Synthesis of the findings from prior reviews (SYN) • Narrative and/or statistical aggregation of prior review findings (AE) 	<ul style="list-style-type: none"> • Identification of areas where more research is needed (RG)

^aGoals based on Rowe (2014)^bBased on Paré et al. (2015), who distinguish and illustrate the review types based on nine dimensions^cSYN synthesizing, AE aggregating evidence, CRI criticizing, TB theory building, RG identifying research gaps, RA developing a research agenda

Table 2 Knowledge development activities in LRs (based on Schryen et al. (2020), pp. 138ff)

Knowledge development activity	Key characteristics
Synthesizing research (SYN)	<ul style="list-style-type: none"> • Mandatory activity in any LR • Summarizes and organizes published knowledge, establishes order in previous research, and makes transparent how research contributions relate to each other • Follows a systematic approach and provides transparency regarding the state and progress of domain knowledge • May take several forms and involve varying degrees of interpretation • May begin by clarifying fundamental aspects such as definitions, domain-relevant variables, relationships between concepts, and domain vocabulary • May reveal central themes and research streams
Aggregating evidence (AE)	<ul style="list-style-type: none"> • Takes theoretical models as a frame, gathers empirical studies, extracts the evidence, and performs statistical aggregation (e.g., meta-analysis or vote counting) to evaluate the degree to which the evidence supports existing theoretical models • Focuses on aggregating effect sizes in relatively homogeneous models and might include qualifications in the form of moderator analyses • Applicable when enough empirical research has accumulated • Meta-analyses are the most common type of review that aggregates empirical evidence
Criticizing (CRI)	<ul style="list-style-type: none"> • Shows that knowledge related to a problem is in some ways inadequate and prevents a domain from progressing • May occur in different forms by problematizing assumptions or identifying methodological, logical, or conceptual problems • In contrast to cumulative extensions of existing knowledge, criticism suggests a revolutionary path that is likely to be irreconcilable with existing knowledge
Theory building (TB)	<ul style="list-style-type: none"> • Provides provisional, possibly conjectural knowledge in the form of new hypotheses and theoretical models that need to be tested by subsequent research • Encompasses developing new theories, and refining or synthesizing theories
Identifying research gaps (RG)	<ul style="list-style-type: none"> • Describes a mismatch between knowledge that is provided by extant research and knowledge that is required or expected • Is expected to stimulate other authors by substantiating a need for research and motivating researchers to close the gaps • Corresponds to the process of spotting gaps in the existing body of knowledge
Developing a research agenda (RA)	<ul style="list-style-type: none"> • Elaborates on how future research should be conducted to achieve meaningful progress, possibly suggesting specific research designs, empirical settings, or offering strategic recommendations • Activity is contingent on the identification of research gaps or a critique of prior research • Develops a vision on behalf of the authors that is oriented towards a promising research goal and a corresponding chart for further research • Should make specific and actionable recommendations that can even take the form of a detailed deployment plan, which could include specific research propositions, suggestions on research designs, and empirical methods

GenAI in the research process

The potential of GenAI to revolutionize academic research extends beyond its capacity to enhance academic writing, such as with tools like Grammarly. Despite its inherent limitations in developing knowledge, it presents a complex set of opportunities and challenges, altering the research landscape regarding how knowledge is created, shared, and consumed (Benbya et al., 2024), together with the evolving role of authors, reviewers, and editors (Yoo, 2024). On one hand, GenAI promises to enhance the efficiency of knowledge synthesis, democratize access to expertise, and streamline the peer review process, thus potentially expediting the knowledge discovery process and mitigating the reproducibility crisis (Alavi et al., 2024). It also offers the prospect of augmenting human capabilities in generating explicit knowledge

from tacit understanding and providing tailored coaching, thus facilitating “long jumps” in the knowledge exploration process (Schwartz & Te’eni, 2024; Alavi et al., 2024; Yoo, 2024). On the other hand, this technological leap is not without its pitfalls. GenAI introduces risks such as biases, ethical concerns, and the potential for hallucination, which could compromise the quality, transparency, and explainability of research outcomes (e.g., Else, 2023; Kankanhalli, 2024; Lund et al., 2023; Ngwenyama & Rowe, 2024). Moreover, there is a looming fear of homogenization in research, undermining innovation, and impacting the norms of scientific discourse (Ngwenyama & Rowe, 2024; Weber, 2024). The limitations of GenAI in effectively identifying gaps in (inter)disciplinary knowledge and conforming to scientific norms further underscore the need for more research to mitigate these challenges. Thus, while GenAI brings forth new

avenues for advancing academic research, it also necessitates a cautious approach to address its inherent risks and ensure that it supports, rather than undermines, epistemic values.

Among those emerging issues, the use of GenAI for conducting LR has garnered particular attention among scholars (e.g., Dasborough, 2023; Dwivedi et al., 2023; Pan et al., 2023). GenAI helps with literature reviews by processing diverse unstructured and structured data to uncover hidden patterns, relationships, and insights within scientific literature (Alavi, 2024). It can retrieve mainstream or dominant views from existing literature, allowing researchers to review, criticize, validate, and extend the baseline with their thought experiments (Ngwenyama & Rowe, 2024). Researchers have effectively used GenAI to support a range of LR activities, such as generating references, analyzing literature, drafting papers, understanding different perspectives, and providing a foundation for theorizing (Jarvenpaa & Klein, 2024). The current application of GenAI in LR shows some initial insights. For example, GenAI has proven its potential to generate effective Boolean queries for systematic literature searches, where it is able to follow complex instructions and generate queries with high precision (Wang et al., 2023). Nevertheless, in another research context, using GenAI for literature searches did not generate ideal results. Gwon et al. (2024) compared the performance of ChatGPT and Microsoft Bing AI in conducting literature searches on Peyronie's disease. Their findings showed that out of 1287 studies identified by ChatGPT, only 7 (0.5%) were directly relevant. In contrast, Bing AI identified 48 studies, of which 19 (40%) were relevant, approaching the human benchmark of 24 relevant studies. The inconsistency in findings highlights the varying performance of GenAI tools in executing literature searches.

Beyond literature search, GenAI can also assist in element mapping and coding of relevant publications; the generative aspect of GenAI allows the researcher to reconceptualize the element maps based on their expertise and insights gathered during the process, followed by the formulation of discussion and conclusion (Pan et al., 2023). For these challenging activities, the reliability and consistency of GenAI have been found to be still on par with those of human researchers (Jenko et al., 2024; Maniaci et al., 2024). Regarding frequent concern over fabrication (information that is plausible-sounding but not factually accurate) and errors generated by GenAI in LR, research shows a great improvement from GPT-3.5 to GPT-4; fabrication was found in 55% of GPT-3.5 citations but just 18% of those in GPT-4 (Walters & Wilder, 2023). Embracing recent advancements in GenAI, studies evaluate the capability of both human researchers and GenAI to delineate the socio-technical requirements for using GenAI in LR. These requirements include avoiding the backward nature of data collection, ensuring transparency of parameters and model weights, facilitating iterative

dialogue between GenAI and researchers, selecting GenAI tools that allow for critical interrogation of data, and maintaining awareness of GenAI's influence on the research process (Ngwenyama & Rowe, 2024).

Some early attempts have been made to empirically evaluate the practical utility of GenAI tools in literature reviews: For instance, Si et al. (2024) conducted a large-scale human study with over 100 NLP researchers to assess whether large language models can generate novel research ideas, finding that LLM-generated ideas were judged as more novel than those from human experts, albeit slightly weaker in feasibility. Similarly, de la Torre-López et al. (2024) presented a survey of AI techniques proposed over the past 15 years to assist researchers in conducting systematic analyses of scientific literature, providing a historical perspective on the evolution of AI in literature reviews. Additionally, Gwon et al. (2024) evaluated the performance of ChatGPT and Microsoft Bing AI in conducting literature searches for systematic reviews, suggesting that while these generative AI tools hold promise, they are not yet sufficiently accurate or feasible for real-time evidence generation in medical research. These early attempts have provided initial empirical evidence on the effectiveness of using GenAI in literature reviews. However, and in contrast to research on how non-GenAI can support LR (e.g., van Dinter et al., 2021; Wagner et al., 2022), the literature is still silent on how GenAI can support epistemological activities when compiling a literature review. To address this gap, our study proposes an insightful perspective guiding the epistemological use of GenAI in LR while emphasizing best practices for human-AI collaboration.

Supporting knowledge development activities with generative AI tools

Approaching our research goal of revealing the potential of GenAI tools for compiling IS LR and to develop recommendations, we proceed by using (a) the sample domain of "IS business value"; (b) several GenAI tools, including ChatGPT (model GPT-4), Perplexity (model GPT-3), Bing AI (now Microsoft Copilot) (model GPT-4), and Google Gemini (model 1.0 pro); and (c) sample queries to illustrate our recommendations. However, our recommendations are not specific to this particular domain or the selected AI tools, models, or queries used. Rather, they are based on and driven by the fundamental nature of different knowledge development activities and the general paradigms underlying large language model-based GenAI tools.

We detail and provide examples of how GenAI tools can support researchers in each of the mentioned knowledge development activities in the sense of human-AI collaboration. We entered numerous prompts into the GenAI tools and present selected examples of the outputs produced to show

their potential. However, not all prompts returned promising results: we show an example of the limited ability of current GenAI tools to support various types of reviews, as discussed in the “[Implications for review types](#)” section.

Synthesizing

Synthesizing research (SYN) involves identifying scholarly work and summarizing, comparing, and contrasting it, ideally, in a concept-centric way (Webster & Watson, 2002). The identification of literature is usually done by querying literature databases, scanning tables of contents, etc.; for a methodology, see, for example, the tutorial by Schryen (2015). Complementing and going beyond these tasks, GenAI may be used to identify literature that adopts a specific perspective on a topic of investigation, be it from a particular theoretical or epistemological perspective. This approach allows the identification of literature from a specific perspective and the organization of its presentation. It also fosters the adoption of a multi-view perspective on a topic. Figure 1 shows a ChatGPT query and answer with which IS business value is viewed from the perspective of the IS success model of DeLone and McLean (1992).

It should be noted that ChatGPT partially hallucinates because “service quality” and “net benefits” do not belong to the categories of IS success mentioned by DeLone and McLean (1992) and should be replaced by the categories “individual impact” and “organizational impact.” While, unsurprisingly, a thorough understanding of the original model requires reading the article of DeLone and McLean, the answer is useful for synthesizing IS business value in several other ways. First, it quickly provides a rough understanding of the nature of the suggested IS success model. Second, the inadvertently included categories of “service quality” and “net benefits” provide starting points for a literature search on the role of these concepts in IS business value. For example, “service quality” (as rated by customers) was found to have a positive effect on the “intended use” of operational CRM technology (Hsieh et al., 2011), which highlights a relationship between the concept of “service quality” and the concept of “intended use,” which is related to the category of “use” included in the DeLone and McLean model. Third, researchers can now dive deeper into the literature to further elaborate how IS has contributed to various elements of the IS success model. Succeeding research steps should involve investigating the provided references and developing appropriate follow-up queries.

In order to adopt a complementary perspective on IS business value, a query may look at this topic from the perspective of the process model suggested by Soh and Markus (1995); Fig. 2 shows an example of such a query and the ChatGPT answer.

Beyond providing references, the answer provides an initial overview of the key differences between the two perspectives and includes ideas of how the two perspectives may complement each other and what are the shared insights. For example, both perspectives recognize the importance of considering the impact of IS on organizational performance, which is a multidimensional construct. Thus, further literature analysis can dive deeper into this concept by analyzing what the IS business value literature, including the two models mentioned above, has found on the impact of IS on different dimensions of organizational performance.

As a synthesis might also include clarifying fundamental aspects, such as definitions and relationships between concepts, a query may be an entry point to a body of concept-defining literature (see Fig. 3). However, it should be noted that one of the shortcomings of GenAI is that its training data may not be up to date, but, if it was able to connect to an up-to-date database of scholar articles, results may improve.


Aggregating evidence

The aggregation of evidence (AE) involves analyzing quantitative data by means of quantitative or qualitative approaches. At the qualitative level, evidence aggregation involves a narrative interpretation of quantitative data. Then, GenAI tools may be used in similar ways as when synthesizing research. In contrast, at the quantitative level, aggregating evidence usually includes the statistical aggregation of empirical studies, such as meta-analysis or vote counting, and involves gathering existing studies, appraising the quality of evidence, determining aggregated effect sizes, and testing their significance (Schryen et al., 2020). The nature of these tasks requires any supporting GenAI tools to include statistical methodologies. In contrast to non-generative AI tools (see, for example, Wagner et al. (2022)), purely text-generative AI tools are not capable of fostering studies that aggregate evidence. However, we envision the development of GenAI tools that generate text based on the statistical analysis of a set of empirical studies.


Criticizing

Criticizing (CRI) reveals that knowledge related to a problem prevents a domain from progressing. It can be implemented by, for example, problematizing assumptions or identifying methodological, logical, or conceptual problems. Contrary to work that cumulatively extends existing knowledge, criticism suggests a revolutionary path that is not likely to be reconciled with existing knowledge (Schryen, 2015). The disruptive character of criticizing prior research makes it challenging to exploit GenAI tools to support this type of knowledge creation, as they rely on historical training data and, thus, can be expected to provide results that foster

Fig. 1 ChatGPT query: the perspective of the model of DeLone and McLean (1992)


You

What does the literature know on IT business value from the perspective of the information systems success model of DeLone and McLean? Explain and provide references.


ChatGPT

The Information Systems (IS) Success Model developed by DeLone and McLean is a comprehensive framework used to evaluate the success of information systems within organizations. This model, first introduced in 1992 and later updated in 2003, has been widely adopted in the literature to assess IT business value from various dimensions. The DeLone and McLean model identifies six interrelated dimensions of IS success:

- System Quality:** The measure of the information processing system itself, including usability, reliability, functionality, and performance.
- Information Quality:** The quality of the output of an IS, which includes relevance, accuracy, timeliness, and completeness of the information produced.
- Service Quality:** The quality of the support that system users receive from the IT service department, including responsiveness, assurance, and empathy.
- Use:** The degree to which an information system is used by the individuals for whom it was intended.
- User Satisfaction:** The users' satisfaction with the information system, reflecting their subjective evaluation of the system and its outputs.
- Net Benefits:** The extent to which the information system contributes to the success of individuals, groups, organizations, or even societies, which can include improvements in productivity, efficiency, decision-making quality, and overall organizational effectiveness.


Research using the DeLone and McLean model has provided valuable insights into how information systems contribute to organizational performance and business value. Studies have explored various sectors including healthcare, education, and e-commerce, demonstrating the model's versatility and adaptability to different contexts.

For references, seminal works and significant studies based on this model include:

- **DeLone, W.H., & McLean, E.R. (1992).** Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60-95. This is the original paper where the model was first introduced, laying the foundation for IS success measurement.
- **DeLone, W.H., & McLean, E.R. (2003).** The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, 19(4), 9-30. This paper revisits the model, updates it, and provides further validation of its constructs.
- **Petter, S., DeLone, W., & McLean, E.R. (2008).** Measuring information systems success: models, dimensions, measures, and interrelationships. *European Journal of Information Systems*, 17(3), 236-263. This paper provides a comprehensive review of IS success measurement, including an extensive discussion on the DeLone and McLean model.


The DeLone and McLean IS Success Model serves as a critical tool for understanding and measuring the value of IT investments and their impact on organizational performance. Its continued relevance and application across diverse studies underscore its utility in guiding both academic research ...

Fig. 2 ChatGPT query: the perspective of the theory of Soh and Marcus (1995)



You

What does the literature know on IT business value from the perspective of the process theory of Soh and Marcus? Distinguish this perspective from the perspective of the information systems success model of DeLone and McLean. Provide References!



ChatGPT

The Process Theory of IT Business Value proposed by Soh and Marcus offers a different perspective on understanding the business value of IT, focusing on the processes through which IT investments transform into business value. This theory is grounded in the premise that IT business value does not automatically result from the mere presence of information technology within an organization. Instead, it emerges through a complex interplay of organizational processes that mediate between IT investments and business value outcomes.

According to Soh and Marcus, there are three main processes involved in the creation of IT business value:

1. **IT Conversion Process:** This involves the acquisition and deployment of IT resources, including hardware, software, and human skills. The efficiency and effectiveness of this process determine the quality and utility of the IT infrastructure.
2. **IT Use Process:** This focuses on how well the IT is actually used within the organization. It emphasizes the importance of user engagement, the alignment of IT with business processes, and the extent to which IT is integrated into daily operations.
3. **Competitive Process:** This refers to how IT is leveraged to create competitive advantages, such as through innovation, differentiation, and enhancement of customer services. It involves the strategic use of IT to impact market position and performance.

The Process Theory of IT Business Value distinguishes itself from the DeLone and McLean Information Systems Success Model by emphasizing the intermediary processes that link IT investments to business outcomes, rather than focusing primarily on the outcome measures of IS success. While the DeLone and McLean model provides a framework for evaluating the success of information systems through dimensions such as system quality, information quality, and net benefits, the Process Theory by Soh and Marcus delves into the mechanisms through which IT resources are transformed into organizational performance and competitive advantage.

For references regarding the Process Theory of IT Business Value: ...

cumulative research rather than revolutionary research. Of course, one may expect to get results that point to (already known) research issues formulated in prior research, but we may not expect to identify uncovered and new research issues. Since large language models are trained on existing data, their knowledge contributions are limited to synthesizing and recombining that data in novel ways, rather than generating fundamentally new knowledge. This limits their value to indirectly inspiring researchers, rather than directly extending knowledge frontiers. As a result, using GenAI tools to critically analyze and advance previous research remains challenging.

However, this challenge does not mean that GenAI tools are inappropriate for researchers who aim to criticize prior research and suggest new research paths. Querying such tools may allow further elaboration of already identified research issues by identifying (supporting or contradicting) evidence and literature that addresses these issues. For example, the LR of Lacity et al. (2010) reviews “the effects that different variables have on IT outsourcing decisions, and the authors challenge the common assumption that outsourcing decisions depend on client size or the size of the IT department” (Schryen et al., 2020, p. 138). Querying GenAI tools may target this issue. For example, researchers may use

Fig. 3 ChatGPT query: definitions of key concepts



You

Provide definitions and relationships between the key concepts of IS business value. Provide references!



ChatGPT

The concept of IT business value focuses on the benefits and outcomes that organizations derive from their investments in information technology. Understanding IT business value involves examining several key concepts and their interrelationships, which encompass the resources invested in IT, the processes that transform these resources into organizational performance, and the metrics used to evaluate success. Here are the definitions and relationships between these key concepts:

1. **Information Technology (IT) Investments:** These are the expenditures on hardware, software, network infrastructure, and human resources aimed at developing, implementing, and maintaining IT systems within an organization. IT investments are the starting point for generating business value through the application of technology.
2. **IT Infrastructure:** This refers to the set of physical and virtual resources that support the overall IT environment, including hardware, software, network resources, and services required for the existence, operation, and management of an enterprise IT environment. IT infrastructure enables the delivery of IT services and solutions.
3. **IT Capabilities:** These are the skills, knowledge, and abilities within an organization that enable the effective deployment, management, and use of IT. IT capabilities include technical proficiency, management expertise ...

The relationships between these concepts can be illustrated as follows:

- **IT Investments** lead to the development and enhancement of **IT Infrastructure** and **IT Capabilities**, which are critical for supporting and enabling business processes.
- **IT Infrastructure** and **IT Capabilities** facilitate the execution and innovation of **Business Processes**, making them more efficient, effective, and aligned with strategic goals.
- The improved **Business Processes**, in turn, lead to enhanced **Organizational Performance**, demonstrating the **IT Business Value** through measurable outcomes such as increased profitability, reduced costs, and strategic advantages.

For references and further exploration of IT business value, the following seminal works provide foundational insights:

- **Brynjolfsson, E., & Hitt, L.M. (1996).** Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending. *Management Science*, 42(4), 541-558. This study provides empirical evidence on how IT investments contribute to firm performance.
- **Melville, N., Kraemer, K., & Gurbaxani, V. (2004).** Information Technology and Organizational Performance: An Integrative Model of IT Business Value. *MIS Quarterly*, 28(2), 283-322. This paper presents an integrative model that outlines the mechanisms through which IT creates business value, emphasizing the mediating role of business processes.

...

queries to investigate the effects of various characteristics of the firm and market, such as scale economies, and the nature and size of business, to approach the question of to what extent outsourcing decisions depend on client size or the size of the IT department.

Theory building

Theory building (TB), which refers to developing and/or refining new hypotheses and theoretical models, represents one of the most challenging and creative activities when compiling a review. While we agree with Wagner et al. (2022) that “theory development is a creative and unstructured endeavor” (p. 215), we also see opportunities for GenAI to support the development or elaboration of theories.

Supporting theory development by means of GenAI tools depends, among other factors, on the type of theory to be developed; for example, Gregor (2006) distinguishes five theory types in IS research: analysis, explanation, prediction, explanation and prediction, and design and action. Although it is beyond the scope of this article to provide a comprehensive analysis of how each of these theory types may benefit from using GenAI tools, we provide an example of a basic type of theory, a theory of analysis, which is used to “describe or classify specific dimensions or characteristics of individuals, groups, situations, or events” (Gregor, 2006, p. 623). We draw on the LR of Schryen (2013), who suggests distinguishing firm performance, innovation, and the protection of resources as types of competitive IS business value. In an attempt to use this distinction as a starting point for developing a taxonomy of competitive external IS business value, researchers may aim to dive deeper into these broad categories by querying GenAI tools for operationalizations and extensions. Figure 4 provides an example. Naturally, the answers given by GenAI tools need to be interpreted with caution, but researchers may now use the answer to look up provided references and investigate mentioned items of firm performance, innovation, and the protection of resources. Consecutive queries may be generated for further information on specific items, with the GenAI tools providing a starting point.

A second example is the use of GenAI tools to support the development of an explanatory theory by identifying causal chains provided in the multidisciplinary literature. For example, in the LR on IS business value by Schryen (2013), mentioned above, answering the research question, “How, why, and when do IS assets, IS capabilities, and socio-organizational capabilities influence each other and jointly create internal value?” (p. 156) is envisioned by using a framework that analyses the relationships between business objectives, CSF/KPI, IS assets, IS capabilities, and socio-organizational capabilities (see Fig. 5). Obviously, the concepts and relationships included in the framework are rooted in several disciplines, including organizational,

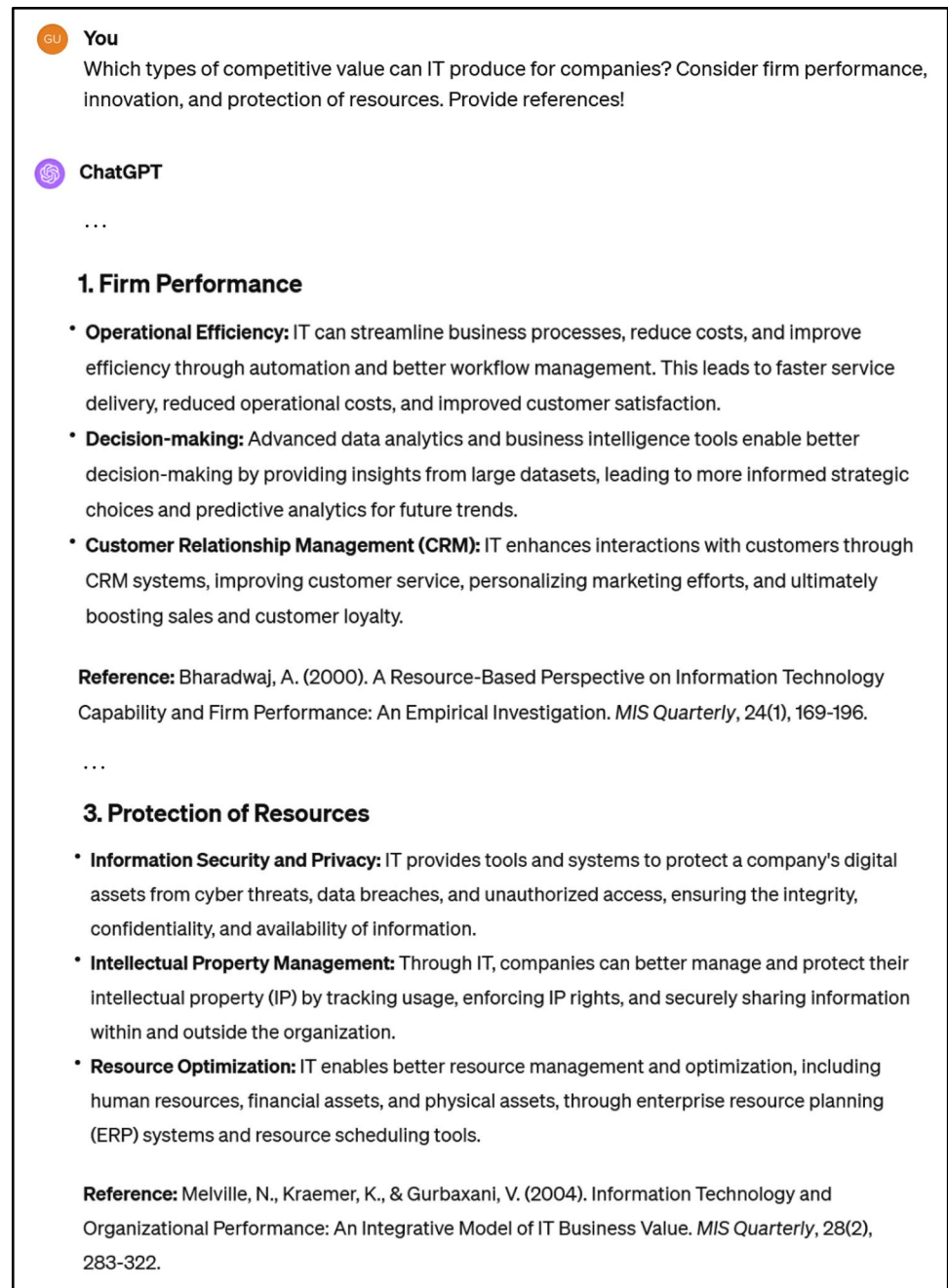
management, and information systems research. Hypothesizing about causal chains thus invites a multidisciplinary perspective and brings together ideas and phenomena from many disciplines. GenAI tools can support this effort by, for example, requesting information about selected parts of the framework, which can either encourage and support hypothesizing about specific causal chains or discourage further analysis. Of course, the information retrieved is unlikely to lead directly to reasonable or innovative hypotheses about causal chains without further (literature or empirical) study, but it may provide useful ideas, concepts, relationships, research directions, etc. that researchers can build on using their knowledge of the IS business value domain and, ideally, their expertise in several of the disciplines involved. Figure 6 shows a sample query and answer focused on the impact of customer relationship management (CRM) processes and CRM systems on improving the “ship-to-target” success factor for companies.

Identifying research gaps

Identifying relevant research gaps (RG) requires an analysis of the existing body of knowledge on a topic, domain, or discipline in terms of relevant research questions and uncharted territories. Although GenAI is well-suited for reproducing existing knowledge, generating truly novel insights or “knowing” the unknown is more challenging and somewhat contrary to their inherent design. As a result, it hardly seems a promising approach to directly query AI tools about such questions and territories, as such tools do not rely on semantic or epistemic models and do not have any topic, domain, or discipline. Rather, it seems quite likely that queries asking for research gaps and questions will be answered by (re)producing text that is composed of research gaps already identified in the literature (Susarla et al., 2023). One effort-saving practice, however, is to utilize GenAI to highlight and summarize the gaps and limitations mentioned by the authors themselves, thereby aggregating these gaps across multiple papers. This approach can potentially uncover patterns and themes that can provide insights to inform future research directions.

A more promising way may be to adopt an indirect approach that aims to identify research gaps by mapping what is known in the literature on a model or theory in the field of investigation. GenAI tools may be queried with a series of questions asking for existing knowledge on particular components of the models and theories and their relationships. The set of results may provide avenues for diving deeper and performing a thorough analysis of potential knowledge gaps by means of a more comprehensive literature search and evaluation process. As an example, we use the model on IS business values suggested by Schryen (2013), who suggests that the competitive value (firm performance, innovation, protection of resources)

Fig. 4 ChatGPT query: categories of competitive IS business value



that a firm may achieve is largely affected by its internal value (IS innovation, socio-organizational capabilities, IS capabilities) and various environmental factors.

Querying GenAI tools may aim to develop entry points for identifying potential uncharted research territories and deriving related research questions through queries that target relationships between internal value, environmental factors, and competitive value (see Figs. 7 and 8).

The queries shown in Figs. 7 and 8 are consecutive queries, with query no. 2 connecting to the answer of query

no. 1 by addressing the challenge of creating a work environment that fosters creativity and innovation. Researchers may now dive deeper into the literature to analyze what is (un)known about this relationship and whether new research gaps may be derived.

Developing a research agenda

The development of a research agenda (RA) requires researchers first to identify research gaps or to criticize prior

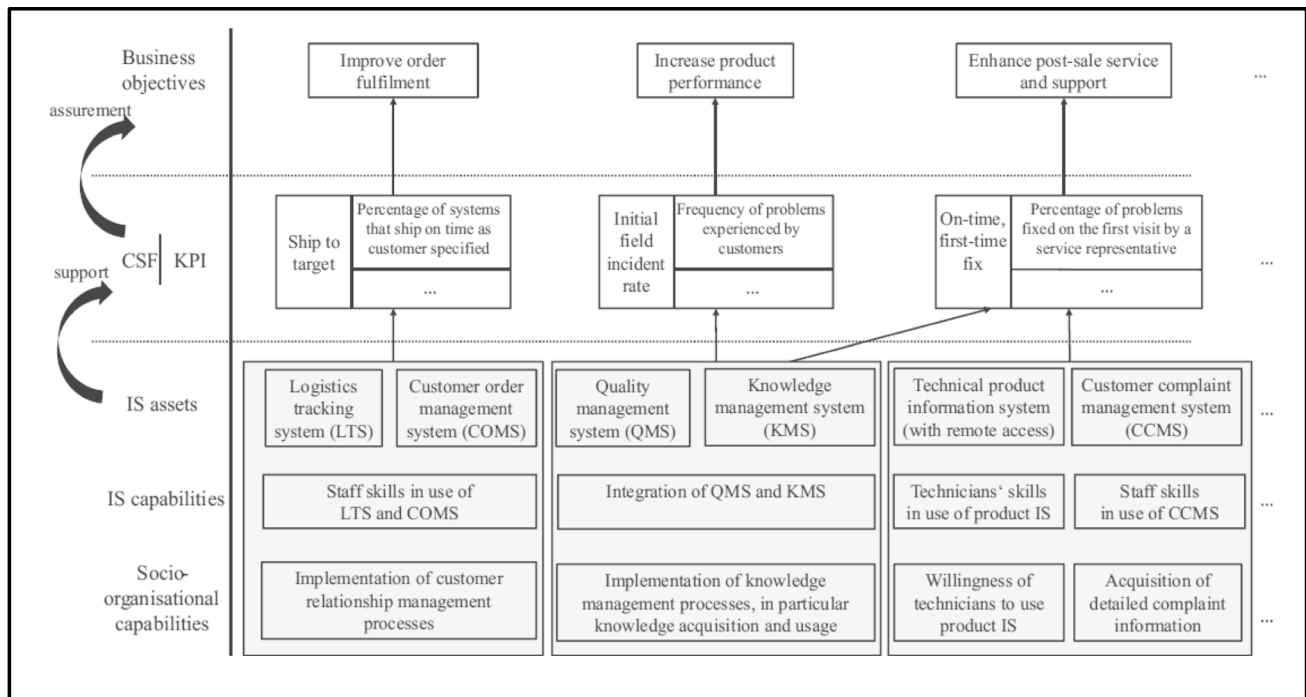


Fig. 5 Research framework for analyzing relationships between business objectives, CSF/KPI, IS assets, IS capabilities, and socio-organisational capabilities (Schryen, 2013, p. 157)

research and then, in a second step, to sketch out a landscape for subsequent research and to make specific and actionable recommendations, which could include specific research propositions, suggestions on research designs, and empirical methods (Rowe, 2014; Schryen et al., 2020).

GenAI tools have been acknowledged to support framing future research (Susarla et al., 2023). However, like the activities of criticizing prior research and identifying research gaps (see the “**Criticizing**” and “**Identifying research gaps**” sections, respectively), developing a meaningful research agenda can hardly be accomplished by simply querying GenAI tools and deriving research propositions, research designs, etc. Yet, answers to a series of queries that dive deeper into research issues, e.g., by looking for (i) insights developed in other academic disciplines; (ii) theories, models, and methodologies used in similar contexts; and/or (iii) research paths that have already been followed but where the approaches adopted should be changed, may inform researchers on potential research paths and inspire them to further investigate those.

We take the LR of Berger et al. (2014) on IS research on “Online Social Networks” (OSN) as an example. The review concludes with a claim, among others, that “[...] future research should especially focus on how to measure organizations’ success in OSN and the related monetary value” and derives a set of research questions, including, “How can the success of OSN activities be measured


beyond reach? Which indicators allow for meaningful statements about the success of OSN activities? How can organizations evaluate the ROI of their OSN activities?” (p. 158). Figures 9, 10, and 11 provide examples of (consecutive) queries that may help researchers transform the sketched research gap into an actionable research agenda.

The three queries shown below are connected to the literature and to each other: query no. 1 uses the research gap identified by Berger et al. (2014) to query what is known on how to measure organizations’ success in OSN and the related monetary value. Queries no. 2 and no. 3 then use two of the resulting potential benefits of using OSN, sales, and marketing activities, as well as improved customer service, to ask how both types of activities can be evaluated in terms of their ROI, which Berger et al. (2014) identified as a key metric in measuring organizations’ success in OSN in terms of monetary value. Researchers may use the answers to all three queries (and further consecutive queries) in their efforts to develop a research agenda on the measurement of organizations’ success in OSN.

Summary of recommendations


The examples in this section show that various knowledge development activities through (standalone) literature reviews can be supported with GenAI tools. Table 3 provides a condensed summary of the capabilities and limitations of GenAI tools, along with recommendations for researchers to

Fig. 6 ChatGPT query: impact of customer relationship management (CRM) processes and CRM systems on improving companies' ship-to-target success factor



You


What is known and unknown about the joint impact of customer relationship management (CRM) processes and CRM systems on improving companies' ship-to-target success factor? Provide references!



ChatGPT

Exploring the joint impact of Customer Relationship Management (CRM) processes and CRM systems on improving companies' ship-to-target success factor encompasses understanding the strategic, operational, and technological dimensions of CRM implementation. This response synthesizes insights from scholarly literature, identifies gaps, and suggests avenues for further research, based on general knowledge up to my last update.

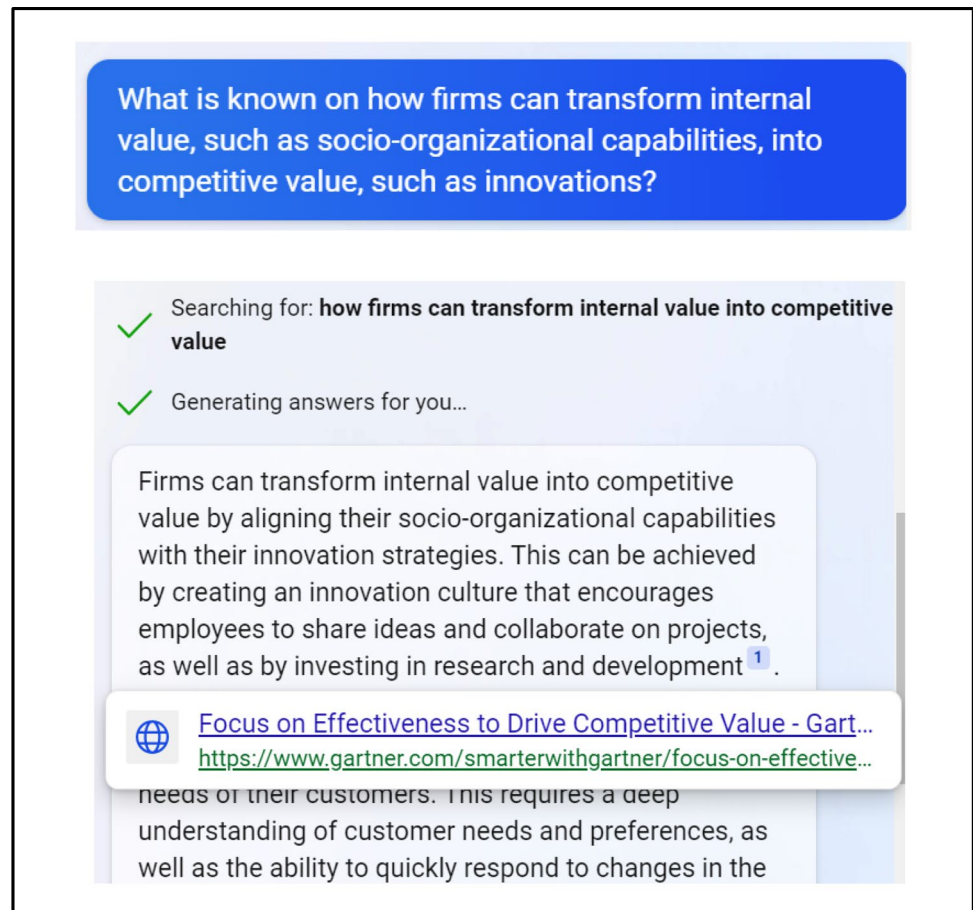
Known Aspects

1. **Strategic Alignment and Customer Focus:** CRM processes and systems are known to enhance strategic alignment towards customer-centric operations, directly influencing a company's ability to meet ship-to-target objectives by improving order accuracy and delivery timelines. References such as Payne and Frow's work on the strategic framework for CRM emphasize this alignment (Payne, A., & Frow, P. (2005). A Strategic Framework for Customer Relationship Management. *Journal of Marketing*). 
2. **Operational Efficiency:** Implementation of CRM systems has been documented to streamline operations, enhance data accuracy, and facilitate seamless communication across departments—key factors in improving ship-to-target metrics. A study by Buttle and Maklan in their book "Customer Relationship Management: Concepts and Technologies" discusses the operational benefits of CRM in logistics and supply chain management (Buttle, F., & Maklan, S. (2019). *Customer Relationship Management: Concepts and Technologies*).
3. **Data Analytics for Predictive Insights:** CRM systems provide businesses with data analytics capabilities that enable predictive insights into customer behavior, demand forecasting, and inventory management, thereby enhancing ship-to-target performance. Research articles in journals like "Decision Support Systems" often explore the use of CRM data for improving logistical operations (*Journal of Decision Support Systems*).

Unknown Aspects

1. **Integration Complexity and Cost-Benefit Analysis:** While the benefits of CRM systems are well-documented, the complexities and costs associated with their integration, especially in SMEs (Small and Medium Enterprises), remain underexplored. The long-term ROI (Return on Investment) and impact on ship-to-target success in varying industry contexts are areas needing further research.
2. **Impact Across Different Industries:** The effectiveness of CRM processes and systems in improving ship-to-target metrics may vary significantly across industries due to differing logistical challenges and customer expectations. Sector-specific studies are less common, presenting a gap in the literature.
3. **Cultural and Organizational Change Management:** The role of organizational culture and change management in the successful adoption of CRM systems and their impact on operational metrics like ship-to-target is an area that is not fully understood. The resistance to change and its mitigation strategies are seldom discussed in the context of CRM implementations.

Fig. 7 Bing AI query no. 1: (missing) knowledge of IS business value creation



use them effectively in the course of knowledge development human-AI collaboration.

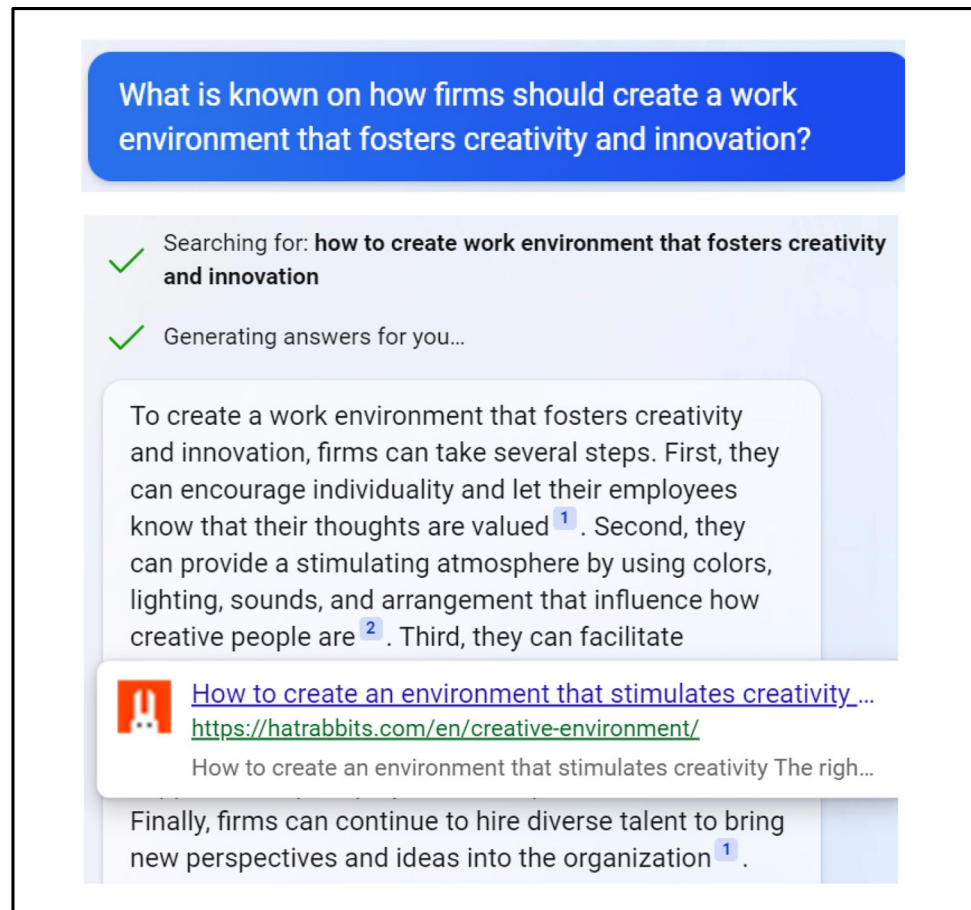
Using GenAI tools does not release researchers from the need to think critically and show creativity. With regard to the former requirement, it holds that, for all activities, it should be taken for granted that the use of GenAI tools requires caution on the part of researchers, as results may include flaws and misleading information, and possibly non-existent references. As with any other tool that supports academic research, results must not be considered “products” that are ready to use in scientific work. The latter requirement includes the challenge to researchers to develop a series of consecutive queries to GenAI tools and to adopt an iterative approach in order to derive promising results. It should be noted that the knowledge development activities considered, which may benefit from using GenAI tools, include both backward-oriented (synthesizing, aggregating evidence) and forward-oriented knowledge development activities (criticizing, theory building, identifying research gaps, developing a research agenda) (Schryen et al., 2020). While it seems hardly surprising that GenAI tools can support backward-oriented knowledge development activities, it may have been considered less obvious that they can also foster forward-oriented knowledge development activities.

Implications for review types

Having analyzed the potential uses of GenAI tools for standalone LRs at the level of knowledge development activities, we proceed with deriving implications for various types of LRs in the IS field (Paré et al., 2015), which can be perceived, from an epistemological perspective, as bundles of knowledge development activities (Schryen et al., 2020); see Table 1 for an overview of the review types. We wish to emphasize that our focus at the level of review types is examining ways of deriving insights into the extent to which key activities of specific reviews can be facilitated through human-AI collaboration. We do not aim to make prescriptive suggestions such as “create a review of type X on topic Y,” as our experiments, similar to those of Susarla et al. (2023), showed discouraging results. We provide two negative examples below when discussing different types of literature reviews.

Narrative, descriptive, and scoping reviews aim to describe phenomena and belong to the review group that primarily summarizes prior knowledge and adopts a broad scope of questions. Narrative reviews are selective, as they do not involve a systematic and comprehensive literature search. These reviews provide a narrative summary of the

Fig. 8 Bing AI query no. 2: (missing) knowledge of IS business value creation



literature and often contribute to identifying research gaps and developing a research agenda. Figure 12 shows a query that asks ChatGPT 4.0 to provide a complete (narrative) literature review on selected IS business value topics. Specifically, the GenAI tool is instructed to produce a literature review that identifies the types of value most extensively discussed in the literature.

While the value types in this example have been discussed in the IS business value literature and indeed represent important areas of research, the review is silent on a large body of research on one of the most extensively studied value types: firm performance (marketing performance, accounting performance) (Schryen, 2013). Overall, the review fails to provide at least a brief overview of the selected subset of the most extensively studied value types as queried. This exemplifies that queries to GenAI tools should not be expected to generate a comprehensive review.

In contrast to narrative reviews, descriptive reviews pursue a representative search strategy. They analyze the extent to which a body of empirical studies in a specific research area supports or reveals interpretable patterns or trends. Beyond summarizing what is known about a topic, they usually also develop recommendations to influence the development of a topic, domain, or method. Scoping reviews adopt

a comprehensive search strategy and examine the extent, range, and nature of research activities. They usually also identify research gaps in the extant literature and develop a research agenda. All these types of reviews can benefit from GenAI tools in organizing literature synthesis, such as clarifying definitions and relationships between concepts and adopting a multi-view perspective. Additionally, GenAI can aid in identifying research gaps and formulating a research agenda through a series of queries, making it suitable for aiding narrative reviews with narrower focuses.

Critical reviews pursue the overarching goal of understanding phenomena and aim to summarize past knowledge and critically analyze the extant literature on a broad topic to reveal weaknesses, contradictions, controversies, or inconsistencies. They often provide a new direction to studies. Due to their focus on criticizing prior research and their disruptive nature, the benefit of GenAI tools for such reviews largely depends on the “creativity” of researchers to use these tools to criticize prior conclusions (see the “Critiquing” section).

Theoretical reviews and realist reviews focus on explaining phenomena. Theory building can occur in different forms, including theory derivation, theory synthesis, and theory analysis. Realist reviews are theory-driven

Fig. 9 Google Gemini query no. 1: identifying types of organizations' success in using OSN

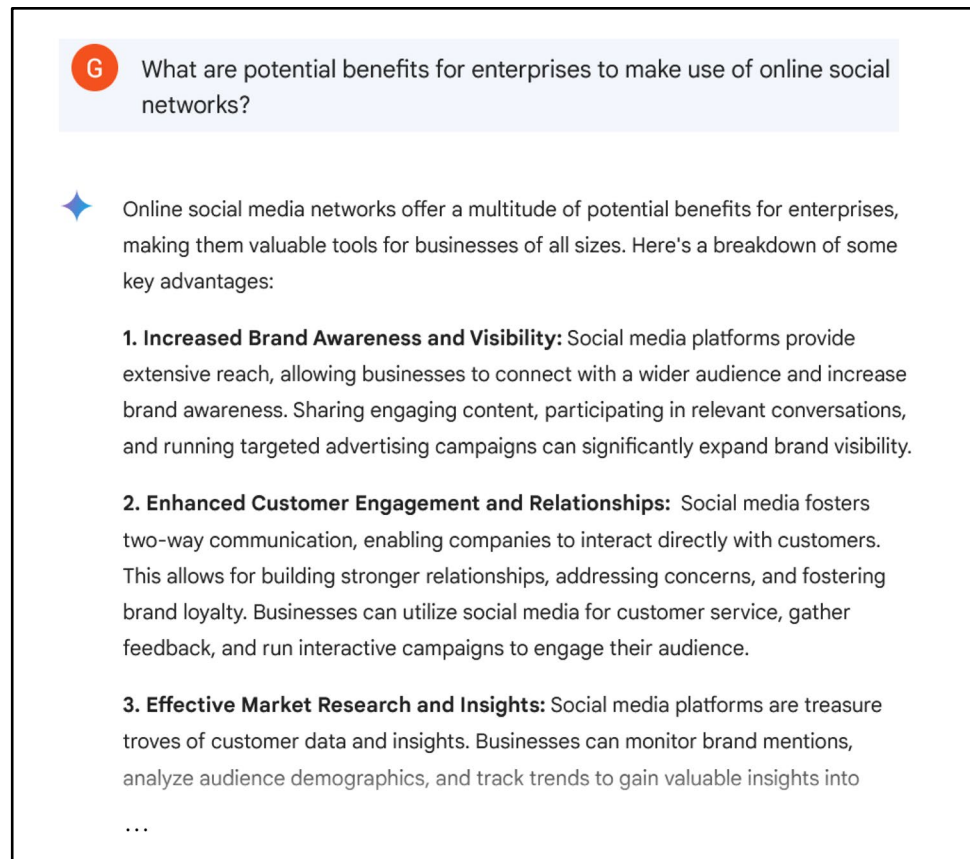


Fig. 10 Google Gemini query no. 2: measuring the ROI of increased brand awareness and visibility

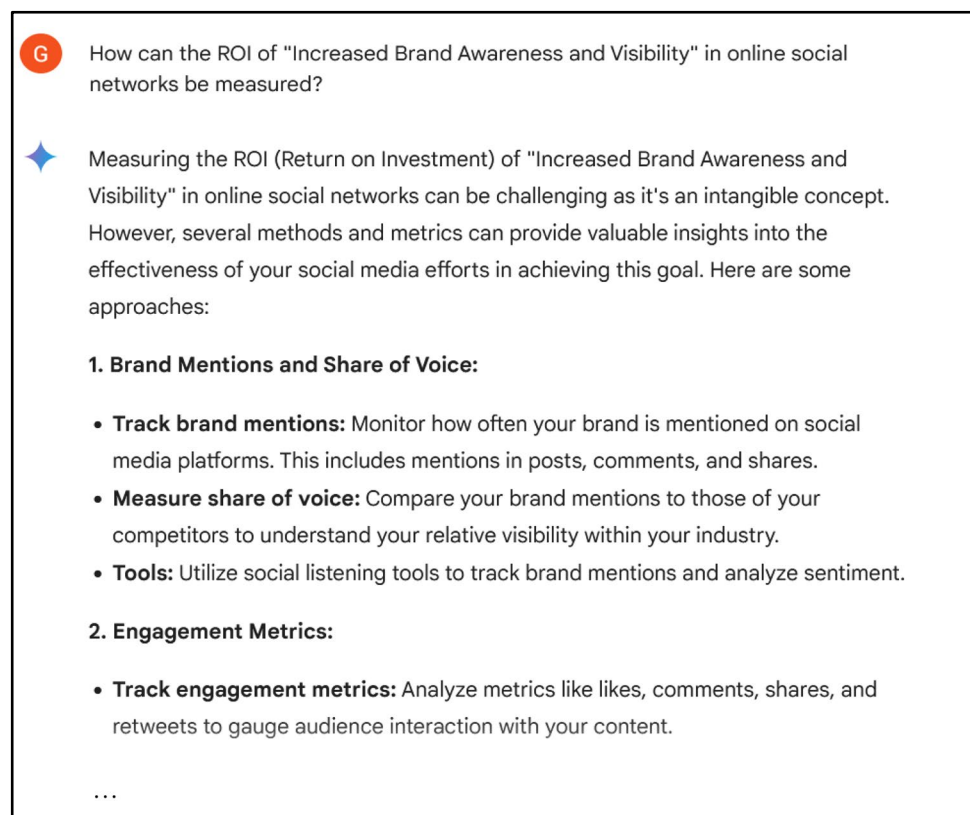



Fig. 11 Google Gemini query no. 3: measuring the ROI of effective market research and insights



How can the ROI of "Effective Market Research and Insights" in online social networks be measured?

✦ Measuring the ROI of "Effective Market Research and Insights" obtained from online social networks, like in the previous response, can be challenging due to the indirect nature of the benefits. However, here are some additional approaches to consider alongside the methods previously mentioned:

1. Social Media Listening ROI Framework:

This framework, developed by social media listening platform Brandwatch, proposes measuring ROI through a combination of:

- **Cost savings:** Quantify the financial resources saved by utilizing social media data instead of traditional market research methods.
- **Impact on business objectives:** Track how social media insights helped achieve specific business goals, such as increased website traffic, improved product development, or better customer service. This can involve assigning a monetary value to the impact based on historical data or industry benchmarks.

...

2. A/B Testing with Market Research Methods:

- Conduct **A/B testing** where one group receives traditional market research and another receives insights from social media.
- Compare the results of both groups on relevant metrics like customer acquisition, product adoption, or marketing campaign performance. This can provide a more direct comparison of the effectiveness of social media insights.

3. Customer Lifetime Value (CLTV):

- Estimate the **CLTV** (Customer Lifetime Value) of customers acquired through marketing campaigns informed by social media insights. This represents the total revenue a customer is expected to generate over their lifespan with your company.
- Compare the CLTV of these customers to the overall average CLTV to assess the potential long-term value derived from social media market research.

...

interpretative reviews; they synthesize evidence and dissemination of findings. GenAI has the potential to assist researchers in theoretical and realist reviews by enhancing the process of theory refinement and development. Through its language generation and analysis capabilities, GenAI can explore and synthesize diverse sources of evidence, facilitating the identification of key patterns, relationships, and explanatory factors. By leveraging GenAI, researchers can efficiently analyze large volumes of literature and extract

insights that contribute to the formulation and refinement of theories, particularly in realist reviews where the goal is to uncover what interventions are effective, for whom, under what conditions, and why.

The final group of reviews, which share the overall goal of data aggregation and integration, consists of meta-analysis, qualitative systematic reviews, and umbrella reviews. They focus on a narrow set of questions. Meta-analysis focuses on the statistical aggregation of evidence. The current

Table 3 Supporting knowledge development activities in LRs with GenAI tools

Knowledge development activity	GenAI capabilities	GenAI limitations	Recommendations for researchers
Synthesizing research	<ul style="list-style-type: none"> Identify diverse literature sources to encourage a multifaceted view on a topic, including the adoption of varying theoretical, epistemological, or methodological perspectives Search literature, complementing traditional manual literature search procedures through databases Concept-centric synthesis through summarization, comparison, and contrastation of identified literature in a concept-centric way, providing a rough understanding of complex models and frameworks 	<ul style="list-style-type: none"> Hallucinations, in particular incorrect concepts that do not belong to established models Outdated training data limit the accuracy and relevance of information and references provided Overemphasize more known literature, overlooking less popular but equally important sources, potentially leading to a biased understanding of the topic 	<ul style="list-style-type: none"> Researchers need to cross-verify GenAI-generated information with original scholarly sources to ensure accuracy Use GenAI in conjunction with up-to-date databases of scholarly articles Critically analyze and refine the insights provided by GenAI, leveraging their own knowledge, experience, and, sometimes, intuition Develop appropriate follow-up queries to further investigate the provided references and deepen the understanding of the identified literature
Aggregating evidence	<ul style="list-style-type: none"> Perform qualitative analysis, such as narrative interpretation of empirical studies based on statistical analysis results 	<ul style="list-style-type: none"> In quantitative analysis, GenAI lacks statistical methodologies for appropriate quantitative aggregation of empirical evidence, such as required for meta-analysis 	<ul style="list-style-type: none"> Researchers may integrate GenAI with statistical tools and consult with GenAI tools to understand and complement the results obtained through quantitative analysis Researchers need to critically evaluate and validate the narrative interpretations generated by GenAI tools against the actual statistical data and analyses
Criticizing	<ul style="list-style-type: none"> Identify and summarize criticisms previously formulated by other researchers Identify contradictions by analyzing a bunch of literature and highlighting the contradictions in the findings and claims 	<ul style="list-style-type: none"> GenAI does not have the capability to formulate “its own” critique on existing literature, which involves questioning whether “things were done right” 	<ul style="list-style-type: none"> Researchers should provide relevant literature sources and formulate queries for GenAI to further elaborate on already identified research issues After GenAI identifies the contradictions in literature, researchers need to conduct high-level problematization to see the root causes of them, such as challenging the underlying assumptions of existing studies
Theory building	<ul style="list-style-type: none"> Support theory development by providing operationalizations and extensions of broad categories, helping researchers develop taxonomies Identify causal chains provided in multidisciplinary literature, supporting the development of explanatory theories by integrating ideas from various disciplines Provide initial insights and references that researchers can use as starting points for deeper investigation into specific items or concepts 	<ul style="list-style-type: none"> The causal relationship proposed by GenAI may be superficial and lack sufficient reasoning GenAI is not capable of generating entirely new theoretical models or hypotheses without human intervention and creativity 	<ul style="list-style-type: none"> With the initial insights provided by GenAI, researchers should leverage their own abstract thinking, creativity, and intuition to theorize the phenomenon of interest Researchers should complement the initial insights provided by GenAI with thorough reviews of related empirical studies to develop well-founded and innovative hypotheses Researchers should employ an iterative process of querying GenAI tools and analyzing the responses to refine and deepen the understanding of specific theoretical constructs and relationships

Table 3 (continued)


Knowledge development activity	GenAI capabilities	GenAI limitations	Recommendations for researchers
Identifying research gaps	<ul style="list-style-type: none"> Highlight and summarize research gaps and limitations previously identified in the literature, potentially uncovering patterns and themes for future research directions Map existing knowledge on particular components of models and theories, aiding in the identification of research gaps 	<ul style="list-style-type: none"> GenAI is inherently designed to reproduce existing knowledge (of research gaps) rather than identify new research gaps, which involves questioning whether “the right things were done” 	<ul style="list-style-type: none"> Researchers should complement the initial gaps provided by GenAI with a thorough literature review through the database to confirm that the identified gaps are indeed underexplored After GenAI provided a summary of knowledge from existing literature, researchers should apply critical thinking to spot the areas that are underexplored
Developing a research agenda	<ul style="list-style-type: none"> Explore insights from other academic disciplines to guide researchers towards potential research paths and inspire further exploration, such as finding theories, models, and methodologies used in similar contexts 	<ul style="list-style-type: none"> Directly querying GenAI for research propositions, research designs, and empirical methods is unlikely to yield actionable research agendas 	<ul style="list-style-type: none"> Based on the insights and gaps identified with the help of GenAI, researchers should consult with literature and apply their own expertise to develop specific and actionable recommendations, including research propositions, designs, and empirical methods, ensuring these are well-founded and contextually relevant

generation of GenAI tools is unable to support such tasks. Qualitative systematic reviews attempt to search, identify, select, appraise, and abstract data from quantitative empirical studies. While employing the typical systematic review process, they use narrative and more subjective (rather than statistical) methods (Paré et al., 2015). Due to its narrow set of research questions, this type of review may particularly benefit from activities that support synthesizing research and identifying research gaps. Umbrella reviews, also referred to as *overview of systematic reviews*, *systematic review of systematic reviews*, and *meta-reviews*, involve various activities of synthesizing prior research, aggregating evidence, and identifying research gaps under a narrower set of research questions. The benefits of using GenAI tools for umbrella reviews are similar to those for qualitative systematic reviews.


As stated above, queries to GenAI tools are not intended to generate complete literature reviews. A further example of an unsuccessful request for a review can be seen in Fig. 13, which shows a request for an “umbrella review” of how the use of artificial intelligence in organizations has positively or negatively affected the organization’s business performance. While acknowledging the difficulty of this task and its lack of ability to provide such an overview, the GenAI tool (ChatGPT 4.0) provides some recommendations for areas of future research and references to explore. However, even these suggestions are quite misleading, as they suggest research in many areas that are not directly related to business performance (e.g., market performance, accounting performance), such as ethical issues and workforce disruption.

Overall, GenAI tools present a fascinating mosaic of potentials within the broad spectrum of literature review types. The realization of these potentials, however, hinges on the critical examination of the capability of GenAI in LR activities and the careful development of the GenAI-researcher collaboration model. Echoing extant literature on using GenAI for LRs (Dwivedi et al., 2023; Jarvenpaa & Klein, 2024; Ngwenyama & Rowe, 2024; Pan et al., 2023), we argue for the imperative for researchers to recognize their unique strengths, such as intuition, nuances discerning, applying deep critical thinking in identifying knowledge gaps, and innovative theorization, as well as GenAI’s relative advantages, such as meta-knowledge base, rapidly scanning vast databases, identifying patterns, and coding themes from excerpts with consistency (Dasborough, 2023; Pan et al., 2023). In addition, researchers must be aware of the limitations of GenAI in conducting LR activities. As we highlighted in our findings, GenAI tends to overemphasize well-known literature while potentially overlooking less popular but equally important sources, leading to a biased understanding of the topic. Struggling with the complexity of critically evaluating prior research, GenAI is not yet suitable for criticizing and uncovering new, unexplored research

Fig. 12 ChatGPT (version 4)
query: a narrative review on the
topic of IT business value


You

Provide a literature review on the value that IT has created for organizations. Distinguish the types of value that have been most extensively discussed in the literature and explain the key findings for each type.


ChatGPT

The impact of Information Technology (IT) on organizations has been a focal point of research within the fields of information systems, management, and organizational studies. IT's value creation for organizations encompasses various dimensions, including operational efficiency, strategic advantage, organizational transformation, and innovation. This literature review highlights the types of value IT has created for organizations and summarizes key findings for each type.

Operational Efficiency

Operational efficiency refers to IT's role in streamlining processes, reducing costs, and enhancing ...

Key Findings:

- IT investments are positively associated with lower operational costs and higher productivity (Brynjolfsson and Hitt, 1996).
- The integration of IT systems across different organizational functions facilitates process efficiency and information flow, leading to better decision-making (Melville, Kraemer, and Gurbaxani, 2004).

Strategic Advantage

Strategic advantage involves using IT to create competitive differentiation and market leadership. ...

Key Findings:

- IT capabilities are critical for developing strategic resources that enable firms to achieve sustainable competitive advantage (Wade and Hulland, 2004).
- Firms that effectively leverage IT for strategic purposes tend to perform better in terms of market share and profitability (Porter and Millar, 1985).

Organizational Transformation

Organizational transformation refers to the profound changes in organizational structure, culture, ...

Key Findings:

- IT is a key enabler of organizational transformation, allowing firms to reconfigure resources and capabilities in response to environmental changes (Venkatraman, 1994).
- The use of IT in transforming organizational structures and cultures is associated with improved organizational performance and innovation (Orlikowski, 1992).

Innovation

Innovation involves using IT to develop new products, services, or business models that create ...


Key Findings:

- IT investments contribute to both incremental and radical innovation, facilitating the development of new products, services, and business models (Yoo, Henfridsson, and Lyytinen, 2010).
- Digital platforms and ecosystems play a crucial role in enabling innovation by allowing firms to leverage external resources and capabilities (Nambisan, Lyytinen, Majchrzak, and Song, 2017).


Conclusion

The literature on IT value creation for organizations highlights multiple dimensions through which ...

Fig. 13 ChatGPT (version 4) query: a narrative review on the topic of IT business value


You

Produce an "umbrella review" of how the use of artificial intelligence in organizations has positively or negatively affected firm performance of the organization. An umbrella review, also called an overview of reviews, is a tertiary type of study that synthesizes relevant evidence from multiple systematic reviews into one accessible and usable document to address a narrow research question. The umbrella review should place particular emphasis on synthesizing findings from previous literature reviews and identifying areas where more research is needed. Provide multiple references that can be used to further explore the literature.


ChatGPT

Creating an umbrella review within this response's scope on the impact of artificial intelligence (AI) on firm performance is a complex task that involves extensive research and analysis of existing systematic reviews and meta-analyses. However, I can guide you through a structured approach on how such an umbrella review might be conceptualized, including potential findings based on existing knowledge and references for further exploration.

Framework for an Umbrella Review on AI and Firm Performance

Objective: To synthesize evidence from multiple systematic reviews on the positive and negative impacts of AI on firm performance.

...

Positive Impacts:

- Efficiency and Productivity:** AI technologies, such as machine learning and process automation, have been shown to significantly increase operational efficiency and employee productivity.
 - Reference for further reading:** Davenport, T.H., & Ronanki, R. (2018). Artificial Intelligence for the Real World. Harvard Business Review.
- Decision Making and Strategic Planning:** AI supports advanced data analytics for better decision-making and strategic planning, contributing to improved competitive advantage and market positioning.
 - Reference for further reading:** Bughin, J., et al. (2017). Artificial Intelligence: The Next Digital Frontier? McKinsey Global Institute.

...

Negative Impacts:

- Implementation Costs and Complexity:** The high costs and complexity of implementing AI systems can be a barrier for some firms, potentially leading to negative ROI in the short term.
 - Reference for further reading:** Huang, M.H., & Rust, R.T. (2018). Artificial Intelligence in Service. Journal of Service Research.
- Workforce Disruption and Ethical Concerns:** AI can lead to workforce disruption, job displacement, and raise ethical concerns related to privacy and surveillance.
 - Reference for further reading:** Manyika, J., et al. (2017). Jobs lost, jobs gained: What the future of work will mean for jobs, skills, and wages. McKinsey Global Institute.

Areas for Further Research:

...

gaps; it also falls short in supporting studies that require evidence aggregation, especially in quantitative research; GenAI's role in theory building and developing research agendas is also limited, as it cannot directly create innovative research propositions without further human investigation and interpretation.

Consequently, we argue for a human-centric synergistic approach where GenAI complements human researchers in LRs. We recommend that the critical responsibilities of reviewing, critiquing, validating theories, identifying gaps, and extending knowledge rest with human researchers. They are poised to make the final decisions on theoretical apparatus selection, ensuring alignment with the research question and contribution to the field. GenAI tools serve not as standalone solutions but as instrumental aides in the hands of adept researchers. Our findings have shown that GenAI tools can automate the retrieval and initial analysis of literature, enriching the LR process by providing a broad, yet surface-level, overview of the existing knowledge landscape. However, their limitations in depth-oriented tasks such as critical evaluation, theory innovation, and knowledge gap are evident. Human researchers are, therefore, envisioned first as the directors and then as overseers to provide guidance, engage with, and refine GenAI outputs to overcome those limitations. This human-centric synergistic approach involves a strategic balance where the efficiency and breadth of GenAI's literature scanning and thematic analysis capabilities are leveraged to set the stage for deeper researcher-led inquiries. Researchers' critical thinking, creative synthesis, and evaluative judgment become the driving forces that navigate and interpret GenAI-generated insights, transforming them into profound contributions to knowledge. By fostering a partnership that capitalizes on the strengths of both GenAI and human researchers, this model aims to advance LR activities in a way that is inclusive, comprehensive, and critical.

Conclusions

In this article, we explore, from an epistemological perspective, how GenAI tools may support IS researchers in developing standalone literature reviews. Our focus is primarily on the innovation goal of GenAI tools in scientific research, as opposed to the communication goal. We argue and illustrate that the effectiveness of GenAI tools largely depends on, and varies based on, specific knowledge development activities. This diversity leads to a more detailed understanding of how GenAI tools can assist in the development of various types of literature reviews, addressing the research question posed in our introduction. While we advise against generating literature reviews in a single step with a single query, our experiments with different GenAI tools lead to

positive results regarding their utility within a human-AI collaborative process. These findings invite further research on how such tools may assist, or possibly hinder, scholars in pursuing the innovation goal of their research.

Recent studies have highlighted significant evolution in the development and usage of these tools within a relatively brief period. This progression includes claims of reducing the capabilities of ChatGPT—often colloquially referred to as “nerfing”—and the continual introduction of innovative features, such as plugins and web browsing capabilities. Therefore, it is crucial for researchers to be well-informed about advancements in this evolving field.

The capabilities and limitations of GenAI, as outlined in our paper, carry significant ethical implications for conducting LR. It is essential that researchers, rather than AI, bear ultimate responsibility for the integrity of their work (Schlagwein & Willcocks, 2023). As GenAI increasingly integrates into LR processes, it is crucial to adhere to the core values of accountability, transparency, replicability, and responsibility (Blau et al., 2024). To maintain the integrity of scientific inquiry, all uses of GenAI in LR must be transparently documented, with explicit attribution distinguishing between human efforts and AI-generated content. This documentation supports the replicability of studies and upholds rigorous scientific standards. Moreover, a primary ethical concern is the management of biases inherent in GenAI tools (Feuerriegel et al., 2024; Stahl & Eke, 2024). Our article highlights that researchers should adopt proactive measures to ensure that biases do not distort research outcomes. These measures include cross-verifying and validating information, supplementing GenAI outputs with up-to-date databases, employing iterative querying, and critically evaluating the outputs of GenAI. Additionally, ensuring the selection of GenAI tools that use diverse training data can help minimize inherent biases. Furthermore, emphasizing human oversight and rigorous critical evaluation remains vital to scrutinize AI-generated outputs effectively. Concludingly, our study enhances the academic community's understanding of GenAI tools' potentials, limitations, and perspectives in supporting researchers in their pursuit of knowledge development through literature reviews.

Despite the promising insights presented in this article, several limitations warrant consideration. First, we acknowledge that our paper relies on hypothetical examples and outputs generated by GenAI tools like ChatGPT, and we recognize that the absence of empirical studies or case studies limits the robustness of our findings. While we have reported on some early attempts in this direction in the background section, our understanding remains incomplete. We encourage further empirical studies to demonstrate the practical utility and limitations of GenAI tools in real-world literature reviews. Such research will enhance the robustness of findings in this field and guide the development of more effective AI-assisted methodologies.

Second, future research should explore the integration of multimodal GenAI tools that go beyond text-based interactions. As GenAI technologies evolve, now incorporating capabilities such as voice recognition and visual data processing, investigating how these tools can be effectively combined with traditional research methodologies would offer valuable insights into how researchers can leverage AI in academic research. In contrast to and complementary to our epistemological focus, this more operational and technical perspective can shed light on how multimodal input can be exploited by GenAI.

Third, the rapid evolution of GenAI tools poses a significant challenge to the generalizability of our conclusions. Findings pertinent today may not hold in the near future as these technologies advance. Another limitation is the inconsistency in tool performance, which refers to the variation in how GenAI tools respond to prompts, both within the same tool by using the same or slightly different queries and using the same queries across different GenAI tools. These differences are due to variations in the underlying algorithms, model architectures, and the training data used for each tool. As a result, the output generated by a GenAI tool can vary significantly depending on the specific prompt or use case or due to the built-in randomness, making it difficult to consistently assess their reliability. Consequently, while the article advocates for a human-centered approach, it should be acknowledged that using GenAI tools in the literature review process must be approached with an understanding of their limitations.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107–136. <https://doi.org/10.2307/3250961>
- Alavi, M., Leidner, D. E., & Mousavi, R. (2024). A knowledge management perspective of generative artificial intelligence. *Journal of the Association for Information Systems*, 25(1), 1–12. <https://doi.org/10.17705/1jais.00859>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/15>. Accessed 24 Oct 2024.
- Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review*, 36(2), 247–271. <https://doi.org/10.5465/amr.2009.0188>
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Bail, C. A. (2024). Can generative AI improve social science? *Proceedings of the National Academy of Sciences*, 121(21), e2314021121. <https://doi.org/10.1073/pnas.2314021121>
- 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(1), 23–36. <https://doi.org/10.17705/1jais.00861>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/13>. Accessed 24 Oct 2024.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the ACM conference on fairness, accountability, and transparency (FAccT '21)* (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Berger, K., Klier, J., Klier, M., & Probst, F. (2014). A review of information systems research on online social networks. *Communications of the Association for Information Systems*, 35(1), 8. <https://doi.org/10.17705/1CAIS.03508>
- Blau, W., Cerf, V. G., Enriquez, J., Francisco, J. S., Gasser, U., Gray, M. L., Jamieson, K. H., Haug, G. H., Hennessy, J. L., Horvitz, E., Kaiser, D. I., London, A. J., Lovell-Badge, R., McNutt, M. K., Minow, M., Mitchell, T. M., Ness, S., Parthasarathy, S., Perlmutter, S., Press, W. H., Wing, J. M., & Witherell, M. (2024). Protecting scientific integrity in an age of generative AI. *Proceedings of the National Academy of Sciences of the United States of America*, 121(22), e2407886121. <https://doi.org/10.1073/pnas.2407886121>
- Buriak, J. M., Akinwande, D., Artzi, N., Brinker, C. J., Burrows, C., Chan, W. C. W., Chen, C., Chen, X., Chhowalla, M., Chi, L., Chueh, W., Crudden, C. M., Di, C. D., Glotzer, S. C., Hersam, M. C., Ho, D., Hu, T. Y., Huang, J., Javey, A., ... Ye, J. (2023). Best practices for using AI when writing scientific manuscripts: Caution, care, and consideration: Creative science depends on it. *ACS Nano*, 17(5), 4091–4093. <https://doi.org/10.1021/acsnano.3c02584>
- Card, N. A. (2011). *Applied meta-analysis for social science research*. The Guilford Press. https://doi.org/10.1111/insr.12011_17
- Dasborough, M. T. (2023). Awe-inspiring advancements in AI: The impact of ChatGPT on the field of organizational behavior. *Journal of Organizational Behavior*, 44(2), 177–179. <https://doi.org/10.1002/job.2695>
- Davison, R. M., Laumer, S., Tarafdar, M., & Wong, L. H. (2023). Pickled eggs: Generative AI as research assistant or co-author? *Information Systems Journal*, 33(5), 989–994. <https://doi.org/10.1111/isj.12491>
- Davison, R. M., Chughtai, H., Nielsen, P., Marabelli, M., Iannacci, F., van Offenbeek, M., Tarafdar, M., Trenz, M., Techatassanasontorn, A. A., Andrade, A. D., & Panteli, N. (2024). The ethics of using generative AI for qualitative data analysis. *Information Systems Journal*. <https://doi.org/10.1111/isj.12504>
- de la Torre-López, J., Ramírez, A., & Romero, J. R. (2023). Artificial intelligence to automate the systematic review of scientific literature. *Computing*, 105, 2171–2194. <https://doi.org/10.1007/s00607-023-01181-x>
- DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60–95. <https://doi.org/10.1287/isre.3.1.60>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan,

- J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Else, H. (2023). Abstracts written by ChatGPT fool scientists. *Nature*, 613(7944), 423. <https://doi.org/10.1038/d41586-023-00056-7>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews*, 1(1), 1–28. <https://doi.org/10.1186/2046-4053-1-28>
- Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, 30(3), 611–642. <https://doi.org/10.2307/25148742>
- Gwon, Y. N., Kim, J. H., Chung, H. S., Jung, E. J., Chun, J., Lee, S., & Shim, S. R. (2024). The use of generative AI for scientific literature searches for systematic reviews: ChatGPT and Microsoft Bing AI performance evaluation. *JMIR Medical Informatics*, 12, e51187. <https://doi.org/10.2196/51187>
- Hart, C. (2009). *Doing a literature review: Releasing the social science research imagination*. Sage. <https://doi.org/10.4236/ajibm.2018.82015>
- Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, 26(2), 38. <https://doi.org/10.1007/s10676-024-09658-1>
- Hsieh, J. P. A., Rai, A., & Xu, S. X. (2011). Extracting business value from IT: A sensemaking perspective of post-adoptive use. *Management Science*, 57(11), 2018–2039. <https://doi.org/10.1287/mnsc.1110.1398>
- Jarvenpaa, S., & Klein, S. (2024). New frontiers in information systems theorizing: Human-gAI collaboration. *Journal of the Association for Information Systems*, 25(1), 110–121. <https://doi.org/10.17705/1jais.00868b>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/6>. Accessed 24 Oct 2024.
- Jenko, N., Ariyaratne, S., Jeys, L., Evans, S., Iyengar, K. P., & Botchu, R. (2024). An evaluation of AI generated literature reviews in musculoskeletal radiology [Article]. *Surgeon*. <https://doi.org/10.1016/j.surge.2023.12.005>
- Kankanhalli, A. (2024). Peer review in the age of generative AI. *Journal of the Association for Information Systems*, 25(1), 76–84. <https://doi.org/10.17705/1jais.00865>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/9>. Accessed 24 Oct 2024.
- King, W. R., & He, J. (2005). Understanding the role and methods of meta-analysis in IS research. *Communications of the Association for Information Systems*, 16(1), 32. <https://doi.org/10.17705/1CAIS.01632>
- Lacity, M. C., Khan, S. A., Yan, A., & Willcocks, L. P. (2010). A review of the IT outsourcing empirical literature and future research directions. *Journal of Information Technology*, 25(4), 395–433. <https://doi.org/10.1057/jit.2010.21>
- Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, 5, 1–9. <https://doi.org/10.1186/1748-5908-5-69>
- Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science*, 9, 181–212. <https://doi.org/10.28945/479>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790. <https://doi.org/10.1016/j.ijme.2023.100790>
- Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shirai, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570–581. <https://doi.org/10.1002/asi.24792>
- Maniaci, A., Saibene, A. M., Calvo-Henriquez, C., Vaira, L., Radulesco, T., Michel, J., Chiesa-Estomba, C., Sowerby, L., Lobo Duro, D., Mayo-Yanez, M., Maza-Solano, J., Lechien, J. R., La Mantia, I., & Cocuzza, S. (2024). Is generative pre-trained transformer artificial intelligence (Chat-GPT) a reliable tool for guidelines synthesis? A preliminary evaluation for biologic CRSwNP therapy [Article]. *European Archives of Oto-Rhino-Laryngology*. <https://doi.org/10.1007/s00405-024-08464-9>
- Majchrzak, A., Wagner, C., & Yates, D. (2013). The impact of shaping on knowledge reuse for organizational improvement with Wikis. *MIS Quarterly*, 37(2), 455–469. <https://doi.org/10.25300/MISQ/2013/37.2.11>
- Mitchell, M., & Krakauer, D. C. (2023). The debate over understanding in AI's large language models. *Proceedings of the National Academy of Sciences*, 120(13), e2215907120. <https://doi.org/10.1073/pnas.2215907120>
- Ngwenyama, O., & Rowe, F. (2024). Should we collaborate with AI to conduct literature reviews? Changing epistemic values in a flattening world. *Journal of the Association for Information Systems*, 25(1), 122–136. <https://doi.org/10.17705/1jais.00869>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/5>. Accessed 24 Oct 2024.
- Paul, D. L. (2006). Collaborative activities in virtual settings: A knowledge management perspective of telemedicine. *Journal of Management Information Systems*, 22(4), 143–176. <https://doi.org/10.2753/MIS0742-1222220406>
- Pan, S. L., Nishant, R., Tuunanen, T., & Nah, F. F. H. (2023). Literature review in the generative AI era – How to make a compelling contribution. *The Journal of Strategic Information Systems*, 32(3), 101788. <https://doi.org/10.1016/j.jsis.2023.101788>
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183–199. <https://doi.org/10.1016/j.im.2014.08.008>
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review-A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, 10(1_suppl), 21–34. <https://doi.org/10.1258/1355819054308530>
- Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing. <https://doi.org/10.1002/9780470754887>
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. Available at <https://openai.com/research/language-unsupervised>. Accessed 24 Oct 2024.
- Ramakrishnan, M., Gregor, S., Shrestha, A., & Soar, J. (2023). Design principles for platform-enabled knowledge commons with an expository instantiation. *Journal of the Association for Information Systems*, 24(5), 1413–1438. <https://doi.org/10.17705/1jais.00824>
- Rivard, S. (2014). Editor's comments: The ions of theory construction. *MIS Quarterly*, 38(2), iii–xiv.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, 52(1), 59–82. <https://doi.org/10.1146/annurev.psych.52.1.59>
- Rowe, F. (2014). What literature review is not: Diversity, boundaries and recommendations. *European Journal of Information Systems*, 23(3), 241–255. <https://doi.org/10.1057/ejis.2014.7>
- Schlagwein, D., & Willcocks, L. (2023). 'ChatGPT et al.': The ethics of using (generative) artificial intelligence in research and science. *Journal of Information Technology*, 38(3), 232–238. <https://doi.org/10.1177/02683962231200411>

- Schryen, G. (2013). Revisiting IS business value research: What we already know, what we still need to know, and how we can get there. *European Journal of Information Systems*, 22(2), 139–169. <https://doi.org/10.1057/ejis.2012.45>
- Schryen, G. (2015). Writing qualitative IS literature reviews—Guidelines for synthesis, interpretation, and guidance of research. *Communications of the Association for Information Systems*, 37(1), 12. <https://doi.org/10.17705/1CAIS.03712>
- Schryen, G., Wagner, G., Benlian, A., & Paré, G. (2020). A knowledge development perspective on literature reviews: Validation of a new typology in the IS field. *Communications of the AIS*, 46. <https://doi.org/10.17705/1CAIS.04607>
- Schwartz, D., & Te'eni, D. (2024). AI for knowledge creation, curation, and consumption in context. *Journal of the Association for Information Systems*, 25(1), 37–47. <https://doi.org/10.17705/1jais.00862>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/12>. Accessed 24 Oct 2024.
- Si, C., Yang, D., & Hashimoto, T. (2024). Can LLMs generate novel research ideas? A large-scale human study with 100+ NLP researchers. arXiv preprint arXiv:2409.04109. <https://doi.org/10.48550/arXiv.2409.04109>
- Soh, C., & Markus, M. L. (1995). How IT creates business value: A process theory synthesis. In *ICIS 1995 Proceedings* (p. 4). <https://aisel.aisnet.org/icis1995/4>
- Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, 74, 102700. <https://doi.org/10.1016/j.ijinfomgt.2023.102700>
- Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. (2023). The Janus effect of generative AI: Charting the path for responsible conduct of scholarly activities in information systems. *Information Systems Research*, 34(2), 399–408. <https://doi.org/10.1287/isre.2023.ed.v34.n2>
- Thomson, D., Russell, K., Becker, L., Klassen, T., & Hartling, L. (2010). The evolution of a new publication type: Steps and challenges of producing overviews of reviews. *Research Synthesis Methods*, 1(3–4), 198–211. <https://doi.org/10.1002/jrsm.30>
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367. <https://doi.org/10.1177/1534484305278283>
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7>
- van Dinter, R., Tekinerdogan, B., & Catal, C. (2021). Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology*, 136, 106589. <https://doi.org/10.1016/j.infsof.2021.106589>
- Visconti, G. (Ed.) (2021). Climate, planetary and evolutionary sciences: A machine-generated literature overview. Springer Nature.
- Wagner, G., Lukyanenko, R., & Paré, G. (2022). Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, 37(2), 209–226. <https://doi.org/10.1177/02683962211048201>
- Walker, L., & Avant, K. (2011). *Strategies for theory construction in nursing* (6th ed.). Pearson.
- Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT [Article]. *Scientific Reports*, 13(1), Article 14045. <https://doi.org/10.1038/s41598-023-41032-5>
- Wang, S., Scells, H., Koopman, B., & Zucco, G. (2023). Can ChatGPT write a good Boolean query for systematic review literature search?. ArXiv. <https://arxiv.org/abs/2302.03495>
- Watson, R. T., & Webster, J. (2020). Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems*, 29(3), 129–147.
- Weber, R. (2024). The Other Reviewer: RoboReviewer. *Journal of the Association for Information Systems*, 25(1), 85–97. <https://doi.org/10.17705/1jais.00866>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/8>. Accessed 24 Oct 2024.
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), xiii–xxiii.
- Yoo, Y. (2024). Evolving epistemic infrastructure: The role of scientific journals in the age of generative AI. *Journal of the Association for Information Systems*, 25(1), 137–144. <https://doi.org/10.17705/1jais.00870>. Available at: <https://aisel.aisnet.org/jais/vol25/iss1/4>. Accessed 24 Oct 2024.

Publisher's N Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.