Information Systems Frontiers
https://doi.org/10.1007/s10796-025-10636-9

The Present and Future of Accountability for AI Systems: A Bibliometric Analysis

Sebastian Clemens Bartsch[1] · Long Hoang Nguyen[2] · Jan-Hendrik Schmidt[1] · Guangyu Du[2] · Martin Adam[3] · Alexander Benlian[1] · Ali Sunyaev[4]

Accepted: 29 July 2025

Abstract

Artificial intelligence (AI) systems, particularly generative AI systems, present numerous opportunities for organizations and society. As AI systems become more powerful, ensuring their safe and ethical use necessitates accountability, requiring actors to explain and justify any unintended behavior and outcomes. Recognizing the significance of accountability for AI systems, research from various research disciplines, including information systems (IS), has started investigating the topic. However, accountability for AI systems appears ambiguous across multiple research disciplines. Therefore, we conduct a bibliometric analysis with 5,809 publications to aggregate and synthesize existing research to better understand accountability for AI systems. Our analysis distinguishes IS research, defined by the Web of Science "Computer Science, Information Systems" category, from related non-IS disciplines. This differentiation highlights IS research's unique socio-technical contribution while ensuring and integrating insights from across the broader academic landscape on accountability for AI systems. Building on these findings, we derive research propositions to lead future research on accountability for AI systems. Finally, we apply these research propositions to the context of generative AI systems and derive a research agenda to guide future research on this emerging topic.

Keywords Artificial intelligence · Accountability · Bibliometric analysis · Research propositions

1 Introduction

AI systems have recently witnessed a rise in diffusion across various applications, including smart digital assistants and medical systems (Adam et al., 2023; Feuerriegel et al., 2024; Jussupow et al., 2022). With an ever-increasing integration into daily lives, ethical considerations gain importance, as scandals involving AI systems highlight substantial risks associated with using AI systems (e.g., Gal et al., 2022; Sunyaev et al., 2025; Zemčík, 2021). Notable examples

✉ Sebastian Clemens Bartsch
bartsch@ise.tu-darmstadt.de

Long Hoang Nguyen
long.nguyen@kit.edu

Jan-Hendrik Schmidt
schmidt@ise.tu-darmstadt.de

Guangyu Du
guangyu.du@kit.edu

Martin Adam
martin.adam@uni-goettingen.de

Alexander Benlian
benlian@ise.tu-darmstadt.de

Ali Sunyaev
sunyaev@tum.de

1 Department of Law and Economics, Chair of Information Systems & E-Services, Technical University of Darmstadt, Hochschulstr. 1, Darmstadt 64289, Germany

2 Department of Economics and Management, Institute of Applied Informatics and Formal Description Methods, Karlsruhe Institute of Technology, Kaiserstr. 89, Karlsruhe 76133, Germany

3 Department of Economics, Chair of Smart Services, University of Göttingen, Platz Der Göttinger Sieben 5, Göttingen 37073, Germany

4 School of Computation, Information and Technology, Technical University of Munich, Bildungscampus 2, Heilbronn 74076, Germany

Published online: 06 September 2025

of AI-induced scandals include racial biases in healthcare, gender biases in recruitment processes, and unlawful arrests due to inaccurate facial recognition (Dastin, 2022; Hill, 2023; Obermeyer et al., 2019). More recently, next-generation AI systems, such as systems building on generative AI, have brought new challenges, including the reinforcement of social biases through harmful recommendations (Gupta et al., 2022; Wan et al., 2023; Zemčík, 2021) and malicious use cases like the creation of deepfakes and the spread of misinformation (Ferrara, 2024). Given the constantly evolving capabilities and risks of such AI systems, accountability has emerged as a foundational principle of AI governance (Novelli et al., 2023; Papagiannidis et al., 2025).

Accountability has been studied for many years in various research disciplines and in different contexts (Hall et al., 2017). However, this appearance has already led to different understandings of accountability (i.e., the obligation to explain and justify one's behavior to others; Bovens, 2007, making it difficult to reach a common understanding (e.g., Kim et al., 2020; Moss, 2023; Wieringa, 2023). For instance, in a political context, accountability relates to responsibility, control, responsiveness, and dialogue, whereas, in a medical context, it relates to a means of ensuring job autonomy (e.g., Batey & Lewis, 1982; Mulgan, 2000). This difference in interpretation of accountability leads to conceptual ambiguity, making it difficult to understand, transfer, and apply findings from different disciplines.

While initial conceptual studies have attempted to reduce the ambiguity surrounding accountability, in particular for AI systems, by offering definitions and frameworks (e.g., Horneber & Laumer, 2023; Novelli et al., 2023; Wieringa, 2020), these studies remain largely theoretical. They typically draw on a limited corpus of literature and do not systematically engage with the broader body of existing research on accountability. In contrast to related areas, such as the study of AI systems (Collins et al., 2021), no comprehensive, literature-based understanding of accountability for AI systems currently exists. This lack of understanding represents a critical gap, given the growing importance of accountability in light of legal regulation (Council of the European Union, 2024), organizational AI governance efforts (IBM, 2022; Microsoft Corporation, 2022), and academic calls for more responsible AI systems (Ågerfalk et al., 2021; Mikalef et al., 2025). To address this gap, we conduct a bibliometric analysis to systematically aggregate and synthesize the fragmented knowledge on accountability for AI systems (Donthu et al., 2021). Such an overview is essential to guide future research and help IS researchers navigate accountability's complex and multi-faceted nature, especially as it relates to emerging AI developments such as generative AI systems. In doing so, we differentiate between research published within the IS discipline and what we define as non-IS research. IS research is primarily characterized by its socio-technical orientation, emphasizing the interplay between algorithmic systems, human actors, and organizational contexts in shaping accountability for AI systems (e.g., Marabelli & Newell, 2023; Sarker et al., 2019). Our study makes a clear operational distinction to situate this perspective within the broader scientific discourse. We define IS research as publications classified within the *Computer Science, Information Systems* category of Web of Science. Consequently, non-IS research encompasses the publications on accountability for AI systems from other disciplines captured by our search. This comparative approach enables us to identify blind spots, shared concerns, and discipline-specific strengths, thereby informing future research directions. Thus, we ask the following research questions (RQ):

> RQ1: What can IS research learn from non-IS research on accountability for AI systems?
>
> RQ2: Where does future research on accountability for AI systems need to go?

To answer these research questions, we conducted a bibliometric analysis, including aggregation and synthesis of 5,809 publications. When answering our first RQ, we distinguished IS from non-IS research disciplines to identify commonalities and differences, thereby revealing three distinct patterns regarding the influence of non-IS research on IS research: First, we observed that IS research often builds on findings from other disciplines rather than drawing on prior work within the IS discipline itself, resulting in a strong reliance on external findings and limited intra-disciplinary referencing. Second, IS research views accountability for AI systems as a virtue, contrasting with non-IS perspectives, which predominantly view it as a governance mechanism. Third, while common ground exists in core ethical considerations, IS research is carving out specialized thematic paths that differ from the broad focus of non-IS research.

Based on these observations, we derived three research propositions for future research to answer our second RQ. Furthermore, the generalized insights we identified provide a valuable foundation for addressing accountability in emerging AI developments, such as generative AI systems. By contextualizing our results within the domain of these next-generation AI systems, we illustrate how broader accountability considerations can be meaningfully applied to this rapidly evolving and increasingly influential class of AI technologies. Therefore, we conclude our study by spotlighting generative AI systems, outlining specific directions for future research on their accountability challenges. In doing so, we derive a research agenda to offer guidance for researchers and practitioners navigating the accountability demands associated with their AI initiatives.

2 Theoretical Background

2.1 Conceptual Understanding of Accountability

Accountability describes a relationship between an actor and a forum in which the actor must justify their behavior to the forum (Bovens, 2007). The forum may question the reasons behind actors' behaviors, so actors need to provide information to justify and explain themselves (Bovens, 2007; Lerner & Tetlock, 1999). For instance, in AI development, accountability translates to a relationship between developers, who assume the role of actors by developing an AI system, and managers, who take the role of a forum. In this relationship, developers (i.e., the actors) must provide managers (i.e., the forum) with justifications and explanations for development decisions. Accountability further enables the forum to impose consequences on actors (i.e., sanctions or rewards), including firing or monetary compensation (Bovens, 2007; Wieringa, 2020).

As a construct, accountability can refer to a virtue or governance mechanism (e.g., Bovens, 2010; Novelli et al., 2023). When referred to as a virtue, accountability describes the positive attitude of individuals toward being responsive to their behaviors (Bovens, 2010). This positive attitude toward being responsive makes individuals more likely to follow and integrate guidelines or rules into their behavior to prevent harm pre-emptively (Beu & Buckley, 2001; Martin, 2019; Novelli et al., 2023). In contrast, accountability as a governance mechanism suggests a more formal and normative approach (Bovens, 2010). It follows the goals of reporting and enforcement by ensuring justifications from actors to forums and providing forums legitimacy to sanction or reward actors (Dubnick, 2005; Novelli et al., 2023).

These two understandings of accountability align with the handling of accountability as being either proactive (i.e., acting before an event occurs) or reactive (i.e., acting after an event occurs) (Novelli et al., 2023). Proactive accountability requires actors to carefully think about behaving in an acceptable and desired manner beforehand. Conversely, reactive accountability requires forums to provide guidelines or rules that actors need to follow and justify themselves when they have not followed them (Novelli et al., 2023). Due to these two distinct understandings, accountability as a construct becomes applicable to diverse contexts and situations to investigate and understand stakeholders' behavior (e.g., Beu & Buckley, 2001; Hall et al., 2017; Wieringa, 2023).

2.2 Accountability Across Different Research Disciplines

Since the conceptual understanding of accountability is broad, accountability faces different uses, depending on the context and research discipline, leading to diverging understandings (e.g., Batey & Lewis, 1982; Mulgan, 2000; Stafford & Stapleton, 2022). For instance, accountability in a political context follows the objectives of responsibility, control, responsiveness, and dialogue of actors against public authorities acting as forums (Bovens, 2007; Mulgan, 2000). Here, accountability predominantly acts as a crucial governance mechanism as it guides and controls individuals to behave in the public's interest, thus preventing their interests from being placed above the public's (Mulgan, 2000). In the case of failure to comply with the public's interest, an actor would face threats of sanctions to ensure and enforce desired behavior (Armstrong, 2005; Bovens, 2007).

In a medical context, accountability helps to guarantee job autonomy and creates, for instance, a direct assignment between healthcare professionals and patients (Batey & Lewis, 1982; Emanuel & Emanuel, 1996). Here, accountability aims to ensure that healthcare professionals are equipped with job autonomy to help patients but do not abuse this autonomy and always act in the interest of patients (Checkland et al., 2004). Thus, accountability motivates healthcare professionals to behave ethically and morally to foster trust in healthcare professionals (Checkland et al., 2004).

IS research usually views accountability in an algorithmic context, referring to it as "the obligation to explain and justify their use, design, and decisions of/concerning the [algorithmic] system and the subsequent effects of that conduct" (Wieringa, 2020, p. 10). Therefore, algorithmic accountability describes a socio-technical relationship between algorithmic systems and their stakeholders, such as developers and users, in which the stakeholders are held accountable for the algorithmic systems (Wieringa, 2020). This attribution of algorithmic accountability toward the stakeholders of an algorithm is crucial since algorithms cannot serve as moral actors and, therefore, cannot be held accountable for their actions and decisions (Cooper et al., 2022).

2.3 Accountability for AI Systems

As accountability for AI systems is becoming increasingly important (e.g., French & Shim, 2024; Li & Goel, 2024; Novelli et al., 2023), this paper centers around accountability for AI systems. Therefore, we first define AI systems by drawing on previous IS research and understand AI systems as algorithmic systems that draw on the three interdependent facets of autonomy, learning, and inscrutability (Berente et al., 2021). Autonomy refers to the ability of AI systems to act independently, often with limited or no human oversight. It implies that humans cannot or partially intervene in AI systems' autonomous decisions (Berente et al., 2021; Citron & Pasquale, 2014). Learning describes the capacity of AI systems to iteratively improve themselves by processing and integrating new data (Berente et al., 2021; Berger et al., 2021). Finally, inscrutability highlights AI systems' inherent complexity, often making their internal operations incomprehensible to humans. Consequently, while humans may understand the theoretical concepts of AI systems, the reasons behind particular outputs and behaviors frequently remain unclear (Berente et al., 2021). Based on these three facets, AI systems attempt to "perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem-solving, decision-making, and even demonstrating creativity" (Rai et al., 2019, p. iii).

The combination of autonomy, learning, and inscrutability makes AI systems powerful tools, leading to widespread adoption across organizations and society (e.g., Adam et al., 2023; Jussupow et al., 2022). However, the same characteristics contribute to AI systems' potential to produce unexpected outcomes, sometimes resulting in significant user harm (Gal et al., 2022; Hill, 2023). For example, contemporary AI systems have frequently exhibited racial or gender bias when providing recommendations (Gupta et al., 2022; Wan et al., 2023). Moreover, generative AI systems introduce further risks of harm, such as deepfakes or misinformation (Ferrara, 2024), amplifying existing societal and ethical challenges when utilizing AI systems.

In response to these growing challenges, researchers, practitioners, and policymakers have emphasized the necessity of establishing AI governance mechanisms. Employing accountability for the design, development, use, and further impacts of AI systems to mitigate risks and address harm has emerged as a cornerstone in such AI governance (High-Level Expert Group on AI, 2019; Thiebes et al., 2021). In the context of AI systems, accountability typically requires stakeholders (e.g., developers, organizations) to explain and justify the outcomes and behaviors of their AI systems to affected parties (e.g., users, regulators, the public) (Horneber & Laumer, 2023; Wieringa, 2020). Therefore, accountability for AI systems often involves transparency (Vössing et al., 2022) or auditability measures (Li & Goel, 2024), ensuring more conscious and responsible behavior throughout the AI lifecycle. For instance, developers might be more inclined to adhere to ethical guidelines and act conservatively and compliantly (e.g., Bartsch et al., 2024; Novelli et al., 2023). Indeed, users of AI systems hold themselves accountable even though the AI systems have failed (Du et al., 2024). In this light, practice recognizes the potential of accountability for AI systems and becomes part of novel regulations (e.g., European AI Act; Council of the European Union, 2024) and ethical guidelines (e.g., IBM, 2022; Microsoft Corporation, 2022; Pichai, 2018).

Despite the apparent potential of accountability for AI systems to mitigate risks, significant challenges exist across practice and research. For instance, practice often struggles with clearly attributing accountability, which faces challenges, such as the many hands problem. Future AI developments, such as generative AI systems, amplify these challenges even more by incorporating more stakeholders (i.e., more hands) and transforming bugs into systemic challenges (Xia et al., 2024). On the other hand, research currently struggles with conceptual ambiguity surrounding accountability for AI systems. One reason for conceptual ambiguity is that accountability for AI systems appears across many disciplines, such as law, computer science, and social science, with different interpretations. This difference in interpretation becomes problematic for inherently multifaceted disciplines, such as IS, which faces difficulty in transferring findings emerging from other disciplines.

Confronted with similar challenges regarding conceptual ambiguity, related research has started developing definitions and conceptual frameworks to combat such conceptual ambiguity (e.g., Aleksovska et al., 2019; Hall et al., 2017; Nguyen et al., 2024; Wieringa, 2020). While these studies have provided valuable depth and nuanced understanding by thoroughly discussing a select number of publications, they mainly relied on interpreting a limited set of papers. Consequently, they may not fully capture the breadth of the discourse or the diverse perspectives across various disciplines engaging with accountability for AI systems. Therefore, the existing body of work requires a broader quantitative overview that aggregates and synthesizes the rapidly expanding research landscape on accountability for AI systems. Such a broad quantitative approach would complement existing insights from in-depth qualitative reviews by providing a systematic map of the field and identifying key research clusters and emerging trends. This approach is especially salient given the rapid advancement of AI systems, especially generative AI systems, which introduce new complexities to accountability. A comprehensive, quantitative synthesis, such as that offered by a bibliometric analysis (Donthu et al., 2021), helps to categorize existing knowledge, identify under-explored areas, and provide robust guidance for future research on accountability for AI systems. Therefore, the transparency provided by a bibliometric analysis enables future research to better navigate the complexities of accountability for AI systems by building upon the collective knowledge across disciplines and addressing neglected perspectives.

3 Methodology

To address our first research question, we conducted a bibliometric analysis to systematically review and synthesize existing research on accountability for AI systems. Based on the insights gained, we developed implications and research propositions to inform future research on AI systems, thereby addressing our second research question. Our study followed an established multi-step bibliometric research process (Donthu et al., 2021; Fosso Wamba & Queiroz, 2023; Kahdan et al., 2022), which included the selection of trustworthy and leading databases with relevant publications, the development of a structured and transparent search protocol, and the subsequent analysis of the retrieved literature guided by key sub-questions (SQs). We selected Web of Science as our primary database due to its extensive coverage of peer-reviewed journals and conference proceedings with high quality in our fields of interest. To construct our search string, we first identified key conceptual terms related to accountability, including *accountability* itself as well as commonly associated constructs such as *responsiveness* (Durante, 2013; Mulgan, 2000; Painter-Morland, 2006), *auditability* (High-Level Expert Group on AI, 2019; Naja et al., 2022; Sanderson et al., 2023), and *liability* (Bovens, 2007; Giesen & Kristen, 2014; Jobin et al., 2019). These terms were then combined with two sets of AI-related keywords. The first subset included technical terms reflecting core AI methods such as *machine learning*, *supervised learning*, *unsupervised learning*, *reinforcement learning*, *deep learning*, and *neural networks*. The second subset focused on application-level terms, including *generative AI*, *generative model*, *large language model*, and *foundation model* (e.g., Fosso Wamba & Queiroz, 2023; Kahdan et al., 2022). The final search string used in our study was as follows:

> *("accountab*" OR "responsive*" OR "liab*" OR "auditability") AND ("AI" OR "artificial intelligence" OR "machine learning" OR "neural network" OR "deep learning" OR "reinforcement learning" OR "supervised learning" OR "unsupervised learning" OR "generative AI" OR "generative artificial intelligence" OR "genAI" OR "generative model*" OR "large language model*" OR "LLM*" OR "foundation model*")*

We used this search string to identify relevant peer-reviewed conferences and journal publications (query date: May 8, 2025). In particular, we first collected titles, abstracts, and keywords from English-written publications until May 2025. Next, we extracted these publications, including metadata, ranging from geographic origin to detailed reference information. The search yielded a total of 5,809 relevant publications. To check the search string's validity and the extracted data's relevance, we manually reviewed a random subset of 300 publications. This manual review confirmed the publications' relevance to accountability for AI systems. Thereafter, we divided the retrieved publications into two distinct datasets. We operationalized our classification using the Web of Science subject categories to distinguish between research inside and outside the IS discipline. Specifically, we leveraged the Web of Science category *Computer Science, Information Systems* to identify the core body of IS literature. For our analysis, we define IS research as any publication classified within this category. This category is curated by Web of Science to group journals focusing on socio-technical topics on the intersection of algorithmic systems, human actors, and organizations. It encompasses many of the leading outlets in the IS research, including the European Journal of Information Systems, Decision Support Systems, Information Systems Frontiers, and Business & Information Systems Engineering, which aligns with the socio-technical perspective central to the IS discipline (e.g., Marabelli & Newell, 2023; Sarker et al., 2019). Correspondingly, we define non-IS research as all other publications retrieved by our search query that are not classified under the *Computer Science, Information Systems* category. In our dataset, this non-IS corpus primarily comprises publications without a socio-technical focus from disciplines such as general computer science (e.g., Nature Machine Intelligence), law (e.g., Computer Law & Security Review), and medicine (e.g., Journal of Medical Internet Research). This operational distinction allows us to systematically compare the discourse on AI accountability within the IS discipline against the broader landscape of contributing fields, revealing unique insights, shared concerns, and potential research gaps. The final IS research dataset encompassed 684 publications, and the non-IS research dataset encompassed 5,125 publications, allowing us to contrast non-IS with IS research. The list of publications can be found in our online appendix (see https://osf.io/36e7x/files/osfstorage).

After identifying and extracting our dataset, we prepared the data for analysis by applying preprocessing steps to the publication abstracts. In line with established bibliometric practices (e.g., Donthu et al., 2021; Kahdan et al., 2022), the abstracts were divided into lemmatized tokens, converted to lowercase, and stripped of stop words, punctuation, and numbers. This preprocessing allowed us to access and analyze each token individually and in relation to others. We implemented the preprocessing pipeline using *Python* and the natural language processing library *spaCy*. Following this step, we conducted a bibliometric analysis that distinguished between contributions from IS and non-IS research. This comparative approach enabled us to identify both shared and divergent patterns, providing a basis for IS research to reflect on and learn from developments in other disciplines. To guide our analysis, we translated our overarching research question into three SQs, drawing on previous bibliometric analyses in the IS discipline (e.g., Fosso Wamba & Queiroz, 2023; Kahdan et al., 2022; Latino et al., 2022; Sahoo et al., 2022; Schöbel et al., 2023). The SQs are as follows.

SQ1: **Which research disciplines** influence the research on accountability for AI systems?

SQ2: **What concepts** are associated with accountability for AI systems?

SQ3: **What topics** does research on accountability for AI systems address?

Our bibliometric analysis used diverse data analysis methodologies to answer the SQs, ranging from descriptive analyses to *Word2Vec* word embeddings (Mikolov et al., 2013). First, we explored **which research disciplines** influence the discourse on accountability for AI systems (**SQ1**). Here, we used the references of each publication and clustered them by their research discipline using their Web of Science categories. With these references, we created a color-coded table indicating referencing patterns. This analysis allows us to untangle the reliance on references from different disciplines and highlight the dependencies between them. Additionally, this analysis enables us to capture and understand whether IS research has formed and established its understanding of accountability for AI systems or whether it predominantly depends on non-IS research.

Second, to answer the question of **what concepts** are associated with accountability for AI systems (**SQ2**), we employed *Word2Vec* word embeddings, transforming 31,738 lemmatized tokens from both datasets into word embeddings with vector length $|\vec{v}|=300$ (Mikolov et al., 2013). We used the t-distributed stochastic neighbor embedding representation from the *scikit-learn* library to simplify the analysis. This analysis allows us to put each token into relation and assign each token an individual meaning. We can subsequently objectify the relations between different terms used in the literature on accountability for AI systems and capture their meaning in the discourse on accountability for AI systems. Thus, word embeddings allow us to search for terms whose semantic meanings are close to accountability for AI systems to analyze relationships between terms and their meanings within the context of accountability for AI systems. By training these word embeddings separately on datasets from non-IS and IS research, we can identify related words to the semantic meaning of accountability for AI systems to highlight and understand differences in the meaning of accountability for AI systems in each research discipline. Therefore, this analysis enhances transparency in understanding the semantic meaning of accountability for AI systems.

Third, by computing co-occurrence patterns, we answered **what topics** research on accountability for AI systems addresses (**SQ3**). We adopted an established binary counting approach and leveraged the *VOSViewer* software to extract and visualize thematic clusters (van Eck & Waltman, 2018). This analysis reveals the structure of key terms within the discourse, highlighting thematic focuses and their connections. By identifying specific topics and their interrelations, we can better understand the current discourse on accountability for AI systems. Furthermore, this analysis helps detect blind spots, particularly when comparing non-IS and IS research, offering valuable insights and guiding future research on accountability for AI systems.

4 Results

4.1 Dependency Analysis – Which Research Disciplines

To answer SQ1, we investigated the cross-disciplinary citation patterns for research on accountability for AI systems, tracing references back to their originating fields (Table 1). Our analysis reveals that besides IS research, key contributions emerge from computer science, engineering, natural sciences, and medicine. Many of these disciplines exhibit strong self-reliance, indicated by high diagonal values. For instance, medicine cites its publications in 55% of cases, natural science in 45%, business in 40%, politics in 35%, engineering in 29%, and computer science in 28% (see diagonals in Table 1). While these disciplines also engage in mutual referencing, their primary focus often remains internal (e.g., psychology cites medicine in 35% of their references, while medicine cites psychology only in 6%).

In contrast, IS research on accountability for AI systems shows comparably low self-citation (9%). Instead, IS research heavily relies on external findings, drawing references from other disciplines, including computer science (25%), medicine (17%), engineering (13%), and natural sciences (11%). This extensive referencing of non-IS publications indicates a strong dependency of IS research in this area on concepts and findings from other research disciplines. Conversely, the influence of IS research is currently limited, as non-IS disciplines do not extensively cite

Table 1. Interdisciplinary Dependency of IS- and Non-IS-Publications

	Comp. Science	IS	Engi-neering	Busi-ness	Natural Science	Geogra-phy	Medi-cine	Psy-chology	Social Sciences	Poli-tics	Ethics	Cul-ture	Others
Comp. Science	28%	25%	15%	12%	8%	16%	5%	5%	14%	13%	12%	17%	11%
IS	7%	9%	3%	4%	2%	3%	3%	2%	4%	2%	5%	3%	3%
Engi-neering	17%	13%	29%	6%	17%	11%	4%	2%	4%	3%	2%	28%	9%
Busi-ness	5%	6%	2%	40%	2%	5%	1%	4%	12%	12%	8%	1%	5%
Natural Science	16%	11%	30%	7%	45%	25%	18%	9%	6%	5%	6%	23%	25%
Geogra-phy	4%	3%	4%	3%	4%	25%	1%	0%	2%	2%	1%	4%	3%
Medi-cine	9%	17%	9%	4%	16%	5%	55%	35%	13%	6%	24%	9%	30%
Psy-chology	3%	5%	1%	6%	1%	1%	6%	32%	8%	4%	6%	2%	5%
Social Sciences	4%	5%	1%	9%	1%	3%	2%	6%	24%	10%	15%	2%	3%
Politics	2%	1%	0%	4%	0%	2%	0%	0%	7%	35%	6%	1%	1%
Ethics	1%	1%	0%	1%	0%	0%	0%	0%	2%	2%	9%	0%	0%
Culture	2%	1%	2%	1%	1%	2%	1%	1%	2%	1%	2%	7%	1%
Others	2%	2%	2%	2%	3%	3%	4%	3%	3%	3%	4%	2%	4%

Red cells indicate that a discipline is highly dependent on another discipline, while blue cells indicate low dependency. Please interpret the table as follows: The value in a column indicates how many percent of the references originate from the respective disciplines (rows). For example, 17% of the references in IS publications (horizontal direction) originate from medicine (vertical direction)

IS research on accountability for AI systems. In contrast, citations to IS research predominantly come from within IS research itself (9%), with computer science (7%) and Ethics (5%) being the most notable disciplines that cite IS research on accountability for AI systems. Therefore, while IS research contributes ideas recognized by adjacent disciplines like computer science, it is not yet widely referenced across non-IS research.

4.2 Similarity Analysis – What Concepts

To answer SQ2, we trained *Word2Vec* word embeddings and used them to form clusters around accountability for AI systems to identify semantically similar meanings (Mikolov et al., 2013). We distinguished between the formed clusters based on the semantically closest words for non-IS research (Fig. 1) and for IS research (Fig. 2).

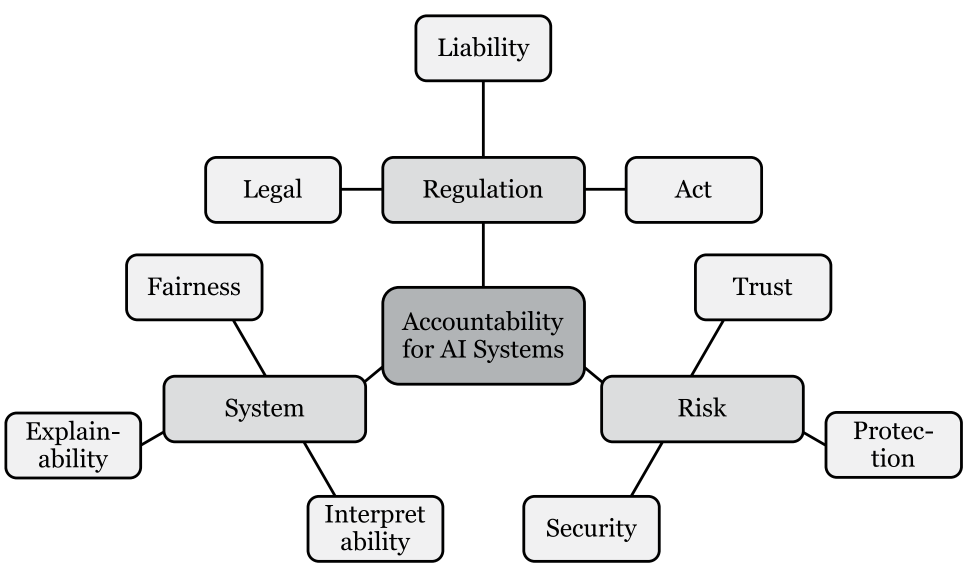


Fig. 1 Formed Clusters Using Word2Vec Representations for Non-IS Research

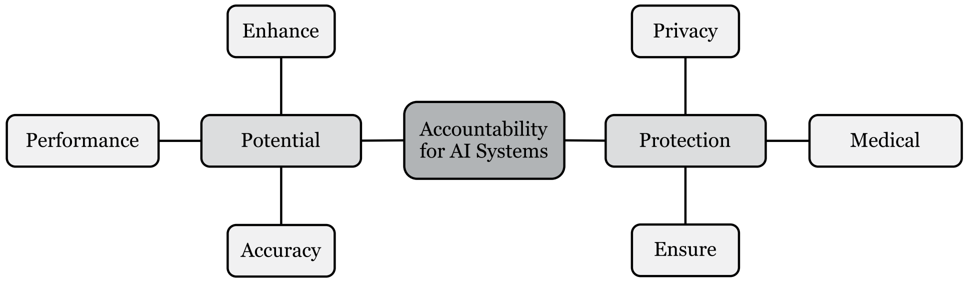


Fig. 2 Formed Clusters Using Word2Vec Representations for IS Research

When focusing on non-IS research (Fig. 1), three key clusters emerge around accountability for AI systems, collectively highlighting facets of AI governance aligning with existing research that understands accountability as a governance mechanism (Bovens, 2010). The first cluster (regulation) underscores the importance of accountability in adhering to formal legal frameworks. This adherence includes compliance with legal requirements, establishing clear liability, and navigating legislative acts. Thus, the first cluster emphasizes how accountability for AI systems ensures that actors operate within established societal and legal boundaries. The second cluster (system) focuses on internal properties and desirable operational characteristics of AI systems. Accountability for AI systems in the second cluster concerns ensuring fairness in outcomes, establishing explainability so that decisions from AI systems are understandable by users, and achieving interpretability of the AI system's internal processes. Therefore, the second cluster highlights that accountability ensures the design and implementation of mechanisms to make AI systems transparent and justifiable. The third cluster (risk) focuses on managing and mitigating potential negative consequences. This risk management process entails building trust in AI systems, implementing robust security measures to prevent misuse or failure, and ensuring adequate protection for individuals or assets. Accordingly, accountability for AI systems involves a proactive approach to identify, assess, and control the risks associated with AI systems to foster their safe and responsible deployment.

In contrast, for IS research (Fig. 2), only two clusters emerge around accountability for AI systems. These two clusters emphasize accountability for AI systems as a virtue (Bovens, 2010) and simultaneously focus on capitalizing on the capabilities of AI systems while considering their safe use.

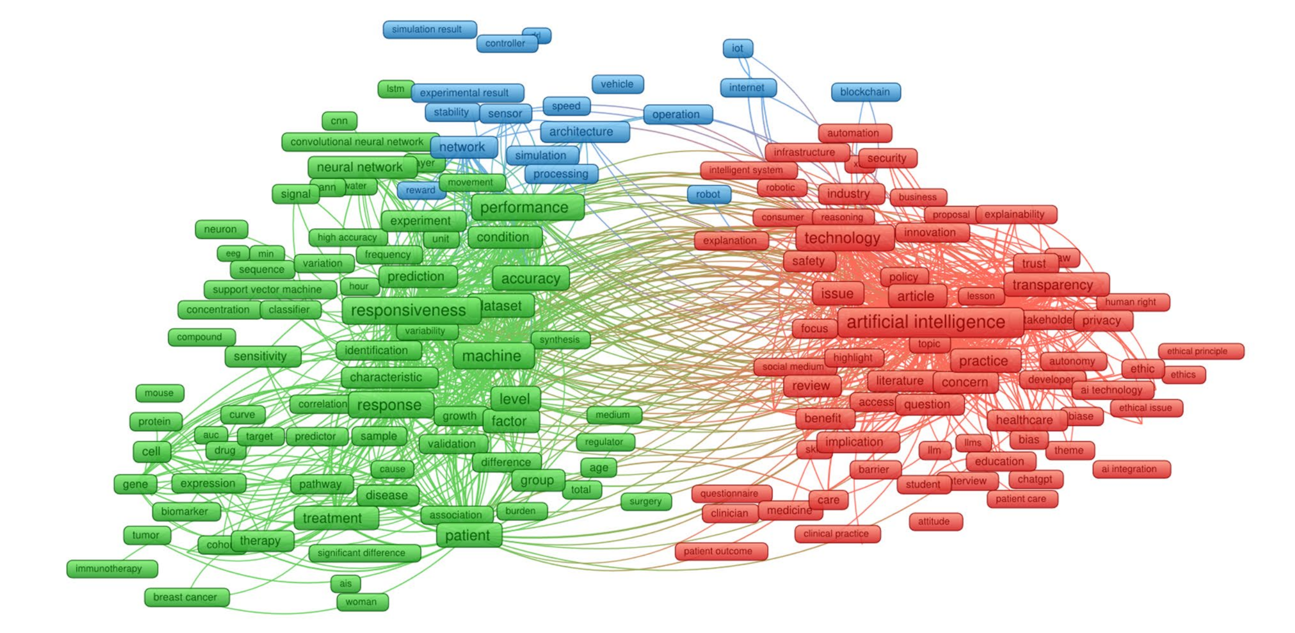
The first cluster (potential) highlights the role of accountability in driving the development and effective use of AI systems. This emphasis on potential involves improving AI system performance, enhancing AI system functionalities and outcomes, and striving for greater accuracy. Thus, the first cluster suggests that accountability for AI systems addresses the need to leverage AI systems' full benefits and advantages. The second cluster (protection) underscores the critical importance of accountability for AI systems in mitigating risks and ensuring responsible AI development. Protection involves privacy guarantees, particularly in sensitive domains like medicine, and ensuring that appropriate safeguards and ethical considerations exist within AI systems. Therefore, this cluster emphasizes that accountability for AI systems also involves a commitment to protect individuals and uphold ethical standards.

4.3 Co-Occurrence Analysis – What Topics

To answer SQ3, we performed a co-occurrence analysis for non-IS and IS research to create thematic clusters (van Eck & Waltman, 2018). These thematic clusters show which topics research focuses on when considering accountability for AI systems. We can better understand the thematic focus in non-IS and IS research by identifying which terms frequently occur in combination. A comparison between research disciplines subsequently uncovers similarities and differences between non-IS and IS research, offering insights into how IS research on accountability for AI systems diverges from non-IS perspectives. Consequently, this analysis not only highlights blind spots in IS research but also provides valuable directions and impulses for future IS research by learning from non-IS research.

4.3.1 Analysis Across Non-IS Research

Figure 3 displays the three thematic clusters derived from the co-occurrence analysis of non-IS research literature on accountability for AI systems. The green cluster emphasizes the applications of AI systems within medical and biological contexts, focusing on machine learning techniques and their performance metrics. The most important terms are *patient*, *treatment*, *machine*, *responsiveness*, *accuracy*, *performance*, *prediction*, and *classification*. This thematic cluster also comprises numerous medical-related terms like *disease*, *gene*, *cell*, *tumor*, *therapy*, *prognosis*, and *breast cancer*, underscoring the emphasis on medical applications. Technical terms such as *support vector machine*, *neural network*, and *deep learning model* are also integral. The central term *machine* connects the green cluster with others, reflecting its broad applicability. Therefore, this cluster highlights



Green: Applications of AI systems within medical and biological contexts

Red: Societal implications, ethical considerations, and governance frameworks of AI systems

Blue: Technical aspects of AI systems, underlying infrastructure, and core AI methodologies

Fig. 3 Cluster Analysis of Non-IS Research

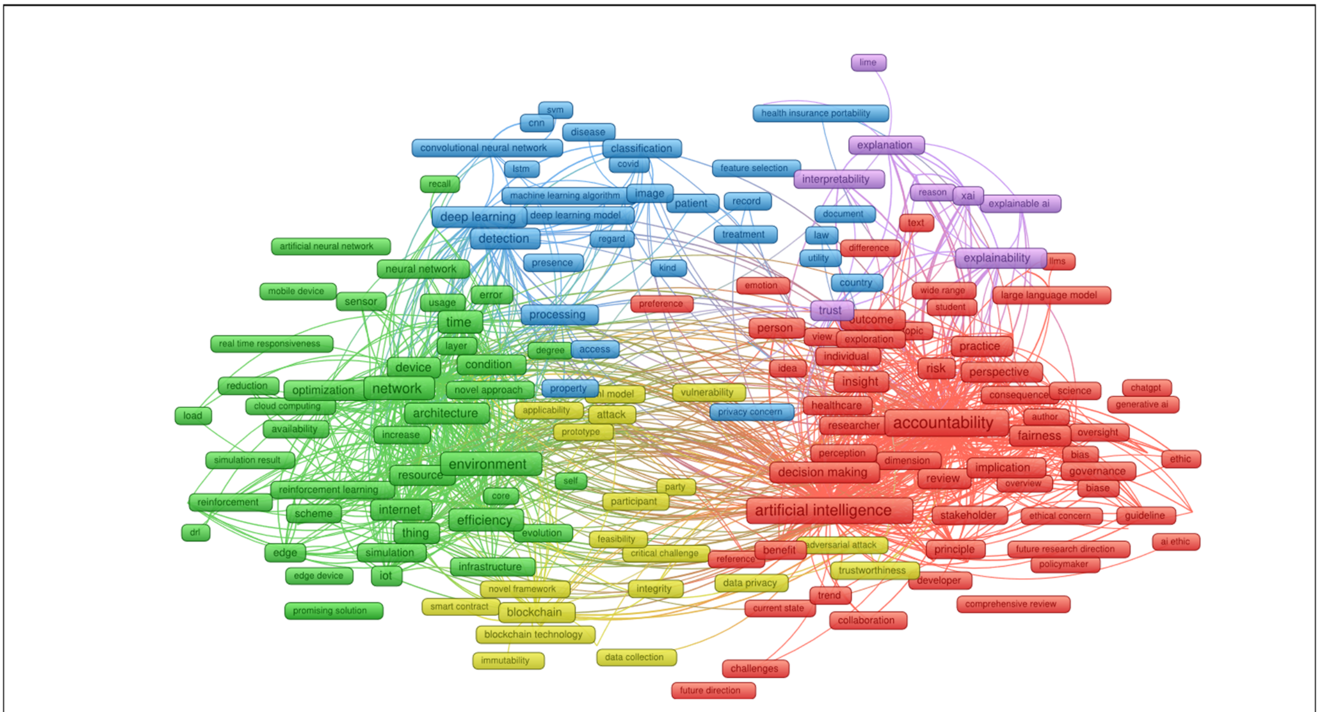
that the discussion of accountability for AI systems often occurs in high-stakes domains, such as medicine, where AI systems' reliability and patient outcomes are critical.

The red cluster focuses on AI-related topics, particularly the broader societal implications, ethical considerations, and governance frameworks. The most important terms are *transparency*, *fairness*, *ethics*, *privacy*, *bias*, *trust*, and *explainability*. This cluster also features terms related to the implementation and impact of AI systems on their stakeholders, such as *issues*, *concerns*, *practices*, and *adoption*. Governance-related terms like *law*, *government*, *policy*, *regulation*, *standard*, and *guideline* are also central to this cluster, indicating their primary importance in the broader discourse on accountability for AI systems. Recent advancements in AI development have also appeared in this cluster through terms such as the *large language model* and *ChatGPT*. The interconnection of the red cluster with other clusters, such as the green and the blue cluster, highlights that the societal and ethical dimensions represented in this cluster are integral to discussions on accountability for AI systems.

The blue cluster focuses on the technical systems, underlying infrastructure, and core AI methodologies associated with AI systems. The most important terms include *network*, *sensor*, *architecture*, *simulation*, *processing*, *operation*, *algorithm*, and *system*. This thematic cluster encompasses various AI methodologies, such as *deep reinforcement learning*, and infrastructural components like *cloud*, *storage*, *internet*, and *IoT* (Internet of Things). Privacy and security considerations are also present through *security* and *blockchain*. This cluster provides the necessary technical foundations and design methodologies for AI systems, and its linkages with the red cluster indicate the integration of these technical considerations with broader ethical and governance concerns.

4.3.2 Analysis Across IS Research

Figure 4 displays the four thematic clusters derived from IS research in the literature on accountability for AI systems. The red cluster concentrates on AI governance, ethical considerations, and societal impact. *Accountability*, *transparency*, *fairness*, *ethics*, *responsibility*, *governance*, *risk*, *bias*, and *trust* are the most important terms. This cluster also includes terms related to *development*, *stakeholder*, *practice*, *implication*, *perception*, *human*, and *society*. The presence of terms like *principle*, *guideline*, and *policy maker* alongside newer concepts such as *large language model (LLM)*, *ChatGPT*, and *generative AI* indicates a comprehensive focus on managing the implications of AI systems. This cluster emphasizes crucial aspects that require active consideration and management throughout the AI lifecycle,



Red: Overarching AI governance, ethics, and societal impact

Green: AI system design, infrastructure, and operational environments

Blue: AI development methods and application domains

Purple: Explainability and interpretability of AI systems

Yellow: Trust-enhancing AI technologies and data governance mechanisms

Fig. 4 Cluster Analysis of IS Research

providing a broad perspective on accountability for AI systems.

The green cluster explores concepts for designing and developing AI systems, their underlying infrastructure, and operational environments. The most important terms are *environment*, *network*, *architecture*, *device*, *internet*, *IoT*, *edge*, *cloud computing*, *efficiency*, and *resource*. AI methodologies such as *reinforcement learning*, *federated learning*, and concepts like *simulation* and *optimization* are also prominent. This cluster represents a general approach to the technical foundations of AI systems development, fostering connections with other thematic clusters and underscoring the necessity of understanding the environment and infrastructure in which AI systems operate.

The blue cluster focuses on specific AI development methodologies and their application, particularly in domains requiring high accuracy and pattern recognition. The key terms are *deep learning*, *detection*, *classification*, *image*, *convolutional neural network*, *machine learning algorithm*, *LSTM* (long short-term memory), and *artificial neural network*. Terms like *recall* and *error* highlight a concern with model performance, and terms like *patient*, *treatment*, and *disease* suggest applications in healthcare. This cluster draws on specific AI development methodologies relevant to the discourse of accountability for AI systems, particularly where model performance and reliability are paramount.

The purple cluster, while smaller, has a distinct focus on the explainability and interpretability of AI systems. The most important terms are *explainable AI*, *interpretability*, *explanation*, *LIME* (Local Interpretable Model-Agnostic Explanations), and *reason*. Concepts like *document*, *text*, and *feature* selection further define this theme. This cluster, connected to the technical (blue) and societal (red) clusters, highlights the increasing importance of understanding and articulating how AI systems arrive at their decisions. Ensuring the explainability and interpretability of AI systems is crucial for fostering trust and facilitating the accountability of these complex technologies.

The yellow cluster centers around trust-enhancing AI technologies and data governance mechanisms, emphasizing blockchain and distributed ledger technology principles. Key terms include *blockchain*, *data privacy*, *integrity*, *security*, *vulnerability*, *attack*, and *data collection*. Concepts like *blockchain technology*, *immutability*, and *smart contracts* further define this theme, suggesting a focus on secure, verifiable, and transparent data handling and system operations. This cluster highlights the critical role of robust technological safeguards and governance frameworks in establishing accountability for AI systems.

4.3.3 Commonalities and Differences Across IS and Non-IS Research

A comparative analysis of the thematic clusters from non-IS and IS research reveals significant commonalities and notable differences in their approaches to accountability for AI systems. Several key commonalities bridge the non-IS and IS research streams. First, both research streams strongly emphasize AI systems' ethical, societal, and governance dimensions. This commonality becomes evident in the red cluster in non-IS research (societal implications, ethical considerations, and governance frameworks of AI systems) and the corresponding red cluster in IS research (overarching AI governance, ethical considerations, and societal impact). Both red clusters share terminology related to *ethics*, *transparency*, *fairness*, *bias*, *trust*, and *governance*, and notably incorporate recent AI advancements such as *LLM* and *ChatGPT*, underscoring a universal concern with establishing responsible AI development and deployment practices. Second, there is a mutual recognition of the importance of the technical foundations underpinning AI systems. The blue cluster in non-IS research (technical aspects of AI systems, underlying infrastructure, and core AI methodologies) and the green cluster in IS research (AI system design, infrastructure, and operational environments) address foundational elements like *network*, *architecture*, *cloud*, and *IoT*. This shared focus indicates a common understanding that accountability intertwines with the technical realities of how AI systems are developed and operated. Finally, applications in healthcare emerge as a common point of interest, albeit with varying degrees of emphasis. The green cluster in non-IS research heavily focuses on medical and biological applications, featuring *patient* and *treatment* as central terms. Concurrently, the blue cluster in IS research (AI development methodologies and application domains) includes healthcare-related terms, such as *patient*, *treatment*, and *disease*. This cluster suggests that IS research engages with accountability in this critical domain.

Despite these commonalities, distinct differences in thematic focus highlight the unique research approaches of non-IS and IS research when addressing accountability for AI systems. One key difference lies in the approach to healthcare and AI applications. While both address this research domain, non-IS research, through its green cluster, demonstrates more depth of focus on medical and biological applications. Here, the performance and reliability of AI systems are paramount, reflecting the critical nature of patient outcomes. In contrast, IS research, primarily within its blue cluster, engages with healthcare but approaches AI systems more broadly. IS research encompasses diverse AI techniques potentially applicable across various domains rather than an equally intensive focus on one specific domain. Another difference is the dedicated focus on explainable and interpretable AI within IS research. The purple cluster in IS research primarily focuses on explainable AI, interpretability, and associated techniques (e.g., LIME). In non-IS research, concepts of explainability and interpretability are present. Still, they are integrated within the broader red cluster (societal and ethical themes) rather than forming a distinct, method-oriented cluster. This difference suggests that IS research is actively carving out a specialized area focused on operationalizing transparency by developing specific explainability methods and tools. Furthermore, the emergence of a cluster for trust-enhancing AI technologies and data governance in IS research marks another difference. The yellow cluster in the IS research highlights topics around *blockchain, data privacy, integrity,* and related concepts like *immutability* and *smart contracts*. While non-IS research does include terms such as *blockchain* and *security* within their clusters, IS research elevates these into a distinct thematic area. This difference indicates a growing IS research focus on leveraging specific technological safeguards and data governance frameworks to establish trust and verifiable accountability for AI systems. Last, while non-IS and IS research have technical clusters, their nuances differ. The blue cluster in non-IS research appears to encompass various fundamental AI development methodologies, including terms like deep reinforcement learning. While covering system design and infrastructure, the green cluster in IS research strongly emphasizes the operational environment, such as *IoT*, *edge, cloud computing*, and system design principles, which include AI methodologies like *federated learning*. This difference could indicate that, compared to non-IS research, the entire range of basic AI algorithm types is less directly in focus, but instead, there is a more pronounced interest in how systems are used and interact in specific operational contexts.

5 Discussion

We conducted a bibliometric analysis to aggregate and synthesize findings on accountability for AI systems to complement existing literature reviews to provide future research guidance on the widespread topic of accountability for AI systems (e.g., Hall et al., 2017; Wieringa, 2020). Against this backdrop, we first operationalized our RQ1 (*what can IS research learn from non-IS research on accountability for AI systems?*) into three SQs (which research disciplines, what concepts, what topics), aiming to investigate accountability for AI systems from different research perspectives. Our results show that IS research references work from outside the discipline more frequently than it engages with prior IS studies, indicating a low level of intra-disciplinary integration. This reliance on external influence suggests that IS research has yet to establish an independent understanding of accountability for AI systems (SQ1, which research disciplines). Next, IS research conceptualizes accountability for AI systems primarily as a virtue by focusing on the potential of AI systems as well as user protection, contrasting with non-IS research's view of it as a governance mechanism revolving around regulation, system properties, and risk management (SQ2, what concepts). Finally, while IS research shows thematic overlap with non-IS research on core topics such as AI governance and ethics, it diverges in how these topics are pursued and linked to specific applications and technical approaches (SQ3, what topics).

5.1 Contributions to Research

In answering RQ1, we aggregated and synthesized existing research and delved into the research area of accountability for AI systems. Our findings provide a multifaceted understanding of the current research landscape on accountability for AI systems. They offer clear insights for future research to learn from and position themselves within the broader discourse on accountability for AI systems. These insights from RQ1 enable us to answer RQ2 (*where does future research on accountability for AI systems need to go?*) by deriving and formulating dedicated research propositions (RP) for future research.

5.1.1 The Need for Stronger Disciplinary Foundations

First, our dependency analysis indicates that IS research on accountability for AI systems predominantly relies on literature from outside the IS discipline, particularly from research disciplines such as computer science, while receiving little reciprocal attention from those research disciplines. This asymmetrical pattern reflects low intra-disciplinary integration and suggests that IS research has yet to establish a widely recognized or consolidated foundation of its own on this topic. While interdisciplinary engagement is essential, especially for complex issues like accountability that involve legal, ethical, and moral dimensions (e.g., Bovens, 2007; Hall et al., 2017; Mulgan, 2000), a lack of internal consolidation within research can weaken its ability to develop coherent perspectives and reduce the visibility of its contributions in the broader academic discourse.

To advance the field, it is essential that researchers not only draw on external insights but also actively engage with, critically assess, and build upon prior research. Such a balanced approach fosters the development of theoretically grounded, discipline-specific perspectives that are both internally robust and externally influential. Strengthening this intra-disciplinary foundation enhances the capacity of a research discipline to contribute distinctive and valuable insights to the wider, multi-disciplinary conversation on accountability for AI systems. Accordingly, we propose our first RP:

> ***RP1****: Research disciplines should strive to balance external integration with internal consolidation by critically engaging with their own prior work to develop coherent and distinctive perspectives on accountability for AI systems.*

5.1.2 Conceptual Divergence and Interdisciplinary Knowledge Transfer

Second, our similarity analysis reveals a significant divergence in the conceptual understanding of accountability for AI systems between non-IS and IS research. Non-IS research understands accountability for AI systems as a governance mechanism, focusing on a control-oriented approach through clusters like regulation and risk to ensure legal adherence and mitigate adverse outcomes. In contrast, IS research primarily understands accountability for AI systems as a virtue, emphasizing proactive development and responsible use through clusters like potential and protection to ensure high performance while also considering privacy. Recognizing these different conceptual understandings across non-IS and IS research is crucial, particularly when researchers aim to apply insights from one research discipline to another as part of knowledge transfer. For example, findings drawn from a study where accountability appears as a governance mechanism (e.g., focusing on enforcement after an incident) might not seamlessly translate or apply to study contexts where accountability appears as a virtue (e.g., focusing on proactive measures). Thoroughly understanding these distinctions helps researchers avoid misalignments and inadequate applications of findings. Furthermore, acknowledging these varied perspectives contributes to a better overall understanding of accountability for AI systems, revealing its different facets and considerations. Therefore, researchers should be mindful of these underlying conceptual assumptions when drawing upon findings from other disciplines and consider how to reconcile or productively contrast divergent perspectives. Following this, we propose our second RP:

> ***RP2****: Future research should consider the different conceptualizations of accountability for AI systems (i.e., as a governance mechanism or a virtue) when transferring knowledge across disciplinary boundaries.*

5.1.3 Cultivating Specializations and Ensuring Holistic Perspectives

Third, our co-occurrence analysis indicates that, while sharing common thematic ground through considering AI governance and ethics, non-IS and IS research cultivate distinct specializations in linking these topics to specific AI applications. For instance, our analysis showed that non-IS research strongly emphasizes medical AI applications. In contrast, medical AI applications are only represented as a side topic within IS research, as IS research focuses on a broader set of AI applications and is more focused on technical AI methodologies, such as explainable AI. Such disciplinary specializations are valuable, allowing targeted contributions based on unique strengths.

However, while such specializations advance discrete areas of knowledge, ensuring accountability for AI systems requires integrating these insights into a broader socio-technical perspective. For instance, as emphasized in IS research, the development of explainable AI techniques is essential yet insufficient if disconnected from how diverse stakeholders interpret explanations and engage with accountability processes. A purely technical explanation does not automatically guarantee accountability if it fails to support stakeholder understanding and action. Likewise, the non-IS focus on AI performance in domains such as medicine is vital. Still, it must be embedded within ongoing ethical reflection to maintain trust and uphold responsible decision-making (Bartsch et al., 2025). Specialized contributions should, therefore, not stand in isolation but be situated within a more comprehensive understanding of how technical, ethical, and stakeholder dimensions intersect in accountable AI systems. Following this reasoning, we propose our third RP:

> ***RP3****: Future research should integrate domain-specific and methodological specializations into a broader socio-technical perspective to ensure a more comprehensive understanding of accountability for AI systems.*

5.2 Summary of the Research Propositions for Accountability for AI Systems

Table 2 summarizes the research propositions we derived to guide future research exploring accountability for AI systems across various topics and research directions. Specifically, these propositions highlight the need to encourage disciplines to build upon their understanding to establish distinct perspectives (RP1), underscore the importance of considering different conceptualizations (e.g., as a governance mechanism or as a virtue) during interdisciplinary knowledge transfer (RP2), and advocate for balancing specialized investigations with a holistic understanding of socio-technical and stakeholder implications (RP3).

5.3 Spotlight on Future AI Developments: Accountability for Generative AI Systems

In the preceding sections, we have extensively examined accountability for AI systems. Our analyses have demonstrated that research on accountability for AI systems lacks consideration of cultural and legislative differences and has limited perspectives, different conceptualizations, and different foci. At the same time, AI systems are experiencing rapid technological evolution, most notably with the emergence of generative AI systems (e.g., Feuerriegel et al., 2024; Hughes et al., 2025; Storey et al., 2025). Although our bibliometric analysis included search terms related to generative AI systems, the results reveal that these transformative

Table 2 Summary and Rationale of the Research Propositions for Accountability for AI Systems

No	Research Proposition	Rationale
1	Research disciplines should strive to balance external integration with internal consolidation by critically engaging with their own prior work to develop coherent and distinctive perspectives on accountability for AI systems.	As exemplified by IS research, some disciplines heavily rely on findings from other disciplines without sufficiently consolidating or being referenced for their understanding, hindering the development of distinct and influential disciplinary perspectives on accountability for AI systems.
2	Future research should consider the different conceptualizations of accountability for AI systems (i.e., as a governance mechanism or a virtue) when transferring knowledge across disciplinary boundaries.	The concept of accountability for AI systems is understood differently across disciplines (i.e., as a governance mechanism or as a virtue), which can lead to challenges in knowledge transfer without careful consideration of the different understandings.
3	Future research should integrate domain-specific and methodological specializations into a broader socio-technical perspective to ensure a more comprehensive understanding of accountability for AI systems.	While specialized perspectives on certain topics (e.g., on explainable AI techniques) are valuable, there is a need to ensure that deep dives into specific aspects of accountability for AI systems remain balanced with a holistic understanding of their broader socio-technical impact.

AI technologies remain marginal in the current accountability discourse. In particular, the co-occurrence analysis (see Sect. 4.3) showed that, while topics like ChatGPT and LLMs are emerging, they have not yet been meaningfully integrated into the conversation about accountability for AI systems. This marginalization suggests a disconnect between the current state of AI developments and the research used to govern it. To bridge this gap and ensure that research on accountability for AI systems remains responsive to the important technological progress of generative AI systems, we propose a research agenda (RA) focused on accountability for generative AI systems.

5.3.1 Generative AI Systems as the Future of AI Systems

Generative AI systems represent a distinct class of AI systems designed to autonomously create new content, such as text and software code (French & Shim, 2024). These systems encompass cutting-edge architectures, such as transformers and large LLMs, which can interpret user queries and generate coherent outputs by computing the next-word probability based on extensive training data (Feuerriegel et al., 2024). Therefore, this architectural advancement marks a significant change compared to classical AI systems, which were often limited to narrow tasks such as data classification (e.g., Niu et al., 2024; Samtani et al., 2023). A defining feature of generative AI systems is their ability to facilitate interaction via natural language, thus eliminating the need for users to interface through complex, multi-dimensional numerical vectors (Storey et al., 2025). This dual capability, accepting natural language as input and returning it as output, has opened up generative AI systems to a broad audience and made such systems more accessible and usable (Storey et al., 2025).

These capabilities have catalyzed the rapid integration of generative AI systems into various sectors of society, fundamentally altering how we work (Chan et al., 2025; Hughes et al., 2025; Storey et al., 2025; Zhong & Yayla, 2025). This change is driven by improvements in user interaction and the underlying characteristics of generative AI systems: Generative AI systems can leverage vast datasets to access knowledge, generate creative outputs, and automate time-intensive tasks (Feuerriegel et al., 2024). Consequently, generative AI systems are no longer confined to narrowly defined functions. Instead, they are increasingly applied across many use cases in the real world (Feuerriegel et al., 2024; Storey et al., 2025). Notable examples include their usage and adoption in higher education (e.g., Hughes et al., 2025), software development and design (e.g., French & Shim, 2024; Zhong & Yayla, 2025), and conversational AI applications such as chatbots (e.g., Heinisuo et al., 2025; Nguyen & Elbanna, 2025). As well as in future scenarios such as healthcare diagnoses (e.g., Katz et al., 2024) and hypothesis generation in academic research (e.g., Li et al., 2025), highlighting the potential of generative AI systems to be used in both analytical and creative domains.

Despite these promising applications, generative AI systems also exhibit critical limitations. Their outputs may be factually incorrect (Feuerriegel et al., 2024), opaque, and difficult for users to interpret (Kenthapadi et al., 2023). Furthermore, these systems are often susceptible to bias and discrimination (Feuerriegel et al., 2024; Zemčík, 2021), raise serious privacy and copyright concerns (Feuerriegel et al., 2024; Kenthapadi et al., 2023), and face challenges related to model robustness and cybersecurity (Kenthapadi et al., 2023). These limitations have intensified calls for increased oversight and control and a deeper engagement with the ethical implications of generative AI systems (French & Shim, 2024; Hammerschmidt et al., 2025; Li & Goel, 2024). Among the most pressing demands emerging from this discourse is greater accountability for generative AI systems (French & Shim, 2024; Li & Goel, 2024). Against the background of generative AI systems, accountability is essential to ensure that the immense power of such systems is managed responsibly and ethically (French & Shim, 2024).

In light of these concerns and calls for more research on accountability for generative AI systems, the following section aims to systematically transfer the insights gained from our previous exploration of accountability for AI systems to the specific case of generative AI systems. Therefore, we propose a research agenda for future research on accountability for generative AI to highlight needed research in this rapidly evolving field.

5.3.2 Accountability for Generative AI Systems

Form an understanding of accountability for generative AI systems within your discipline first before spreading them to other disciplines: IS research, while increasingly engaging with the topic of accountability for AI systems, continues to draw heavily from adjacent disciplines such as computer science, with limited internal referencing and disciplinary consolidation. This asymmetry suggests that IS research may lack a sufficiently distinct or influential voice in the broader discourse on accountability for AI systems. Notably, while such an interdisciplinary approach is essential, given that accountability is inherently a multi-faceted concept encompassing technical, ethical, legal, and organizational dimensions (e.g., Bovens, 2007; Mulgan, 2000), other disciplines' lack of reciprocal engagement indicates limited impact and visibility of IS contributions in return.

This issue is critical in the context of generative AI systems, which introduce new challenges that extend traditional disciplinary boundaries. Generative AI systems not only reshape how individuals and organizations interact with technology through natural language interfaces, automation of creative tasks, and large-scale deployment but also raise novel accountability concerns related to explainability, misuse, misinformation, and ethical governance (e.g., Feuerriegel et al., 2024; Kenthapadi et al., 2023). Consequently, robust disciplinary engagement is necessary to advance context-sensitive, actionable, and theoretically grounded frameworks for understanding accountability for generative AI systems. To contribute to such developments, disciplines such as IS research must move beyond relying predominantly on external accountability frameworks. Instead, disciplines should actively build upon, critique, and consolidate insights from within the field to form coherent, cumulative perspectives that address the specific challenges posed by generative AI systems.

Strengthening the disciplinary foundations will strengthen the internal cohesion of the disciplines and increase their external relevance and impact. A more mature and self-sustaining research tradition will allow the discipline to engage in a reciprocal dialogue with neighboring disciplines and contribute meaningfully to the multidisciplinary discourse on accountability for generative AI systems. Strengthening this research foundation is particularly important given the rapid pace of innovation in generative AI, which risks outpacing the development of theoretical and empirical frameworks for governance and accountability. In light of these considerations, we raise the following questions for our research agenda:

> ***RA1.1****: How do the accountability challenges of generative AI systems reshape foundational assumptions of accountability for AI systems within individual disciplines?*

> ***RA1.2****: How can disciplinary perspectives be combined to create a coherent and cumulative understanding of accountability for generative AI systems?*

> ***RA1.3****: How can disciplines engage more effectively in reciprocal interdisciplinary discourse on accountability for generative AI systems once a strong disciplinary foundation is established?*

Addressing RA1.1, researchers could revisit and critique foundational accountability frameworks (e.g., Bovens, 2007; Novelli et al., 2023; Wieringa, 2020) within their discipline to assess their applicability to generative AI systems with emergent, probabilistic, and autonomous behaviors. In this context, researchers might integrate insights from socio-technical systems theory (Appelbaum, 1997; Ogie et al., 2022) or other related theories used within specific disciplines. Addressing RA1.2, researchers might specifically investigate and explore integrative frameworks for accountability for generative AI systems that combine these dimensions across disciplines to reduce the heterogeneity within existing research (e.g., Star & Griesemer, 1989). Finally, addressing RA1.3, researchers can start translating disciplinary insights about accountability for generative AI systems into language and constructs accessible to adjacent disciplines. In this context, comparative studies, interdisciplinary workshops, and collaborative publications could serve as a means for translating disciplinary insights and knowledge integration.

A comprehensive view of understanding accountability for generative AI systems: Depending on the discipline, accountability for AI systems either appears as a governance or a virtue. Since non-IS research often adopts a control-oriented governance perspective, it primarily frames accountability for AI systems as a mechanism to ensure regulatory compliance, mitigate risks, and manage legal liabilities. This perspective typically emphasizes post-hoc enforcement and safeguards for organizations. In contrast, IS research tends to conceptualize accountability more proactively and normatively, treating it as a virtue embedded in the design and development of AI systems (e.g., Schmidt et al., 2025). This perspective centers on fostering trust, promoting ethical awareness, and enabling responsible innovation, often through practices encouraging transparency and privacy protection.

This conceptual divergence becomes particularly consequential in the context of generative AI systems, which amplify the stakes of accountability due to their autonomy, opacity, and global reach. The dual nature of generative AI systems, capable of producing high-impact outputs while operating as black-box systems, demands regulatory oversight and proactive responsibility. As such, transferring knowledge across disciplines without appreciating these conceptual nuances risks producing incomplete, inconsistent, or even incompatible frameworks for accountability for generative AI systems.

For example, applying accountability from a governance-focused perspective to the design of generative AI systems might overlook critical elements of trust and ethical foresight. Conversely, relying solely on virtue-based perspectives in regulatory enforcement could fail to address liability and organizational responsibility when deployed generative AI systems cause harm. Hence, recognizing and bridging these conceptual differences is essential for meaningful interdisciplinary dialogue and for developing integrative accountability frameworks that are operationally viable and ethically robust when designing and developing generative

AI systems. Therefore, we raise the following questions for our research agenda:

> ***RA2.1***: *How do different conceptualizations of accountability for AI systems, such as governance mechanisms or virtues, shape the development, deployment, and regulation of generative AI systems?*

> ***RA2.2***: *How can governance- and virtue-based conceptualizations of accountability for AI systems be integrated to create comprehensive accountability frameworks for managing generative AI systems?*

> ***RA2.3***: *Which accountability mechanisms are most effective in shaping accountability as a governance mechanism or as a virtue to govern generative AI systems?*

Addressing R2.1, researchers could examine how these divergent conceptualizations influence policy formation, organizational practices, and technical design of generative AI systems (e.g., Bartsch & Schmidt, 2023). For example, virtue-oriented approaches may encourage developers to adopt a more value-sensitive design, while governance frameworks may prioritize auditability and enforcement of generative AI systems. Comparative case studies across disciplines may uncover how these orientations manifest and interact in practice. Addressing R2.2, researchers might explore how governance-based accountability frameworks (e.g., AI auditing; Raji et al., 2020 can be effectively aligned with virtue-based accountability frameworks (e.g., proactive accountability approaches; Novelli et al., 2023) when managing generative AI systems throughout their lifecycle. Finally, addressing R2.3, researchers can empirically investigate how design artifacts (e.g., Adam, 2022; Schmidt et al., 2025) can shape accountability for generative AI systems as a governance mechanism or as a virtue. Such an investigation would allow researchers to effectively control accountability for generative AI systems either as a virtue or as a governance mechanism and integrate it into their accountability framework for generative AI systems.

Bringing special considerations of accountability for generative AI systems back into a general and socio-technical perspective: Non-IS and IS researchers commonly engage with topics such as AI governance and ethics, but often apply them to different AI applications. Non-IS research, for instance, demonstrates a strong orientation toward domain-specific AI applications, particularly in medicine. In contrast, IS research concentrates on technical AI methodologies such as explainable AI, reflecting the discipline's emphasis on developing and implementing socio-technical systems. These specializations within disciplines are necessary and valuable: They enable researchers to create deep, context-sensitive insights and contribute in areas aligned with their methodological strengths and theoretical traditions. However, such specialization also presents risks, such as the fragmentation of accountability discourse for AI systems, especially in emerging and rapidly evolving contexts such as generative AI systems.

Generative AI systems differ from many earlier forms of AI systems in that they moved toward general-purpose applications that are deployed across various domains, interacting with diverse stakeholders, and generating unstructured outputs like natural language, code, and images (e.g., Ferrara, 2024; Feuerriegel et al., 2024; Hughes et al., 2025; Pinski & Benlian, 2024; Söllner et al., 2025). As a result, researchers and practitioners who develop specialized approaches to understanding accountability for generative AI systems, whether they focus on technical transparency or domain-specific ethics, may fall short if they do not integrate these approaches into a broader, socio-technical understanding of how society adopts and integrates these systems.

For example, IS research focusing on the technical dimensions of explainable generative AI systems must also consider how different user groups, such as patients, developers, or managers, interpret and act on these explanations in practice. Similarly, non-IS research focusing on domain-specific research, such as medicine, must reflect on how stakeholders build, train, and govern generative AI systems outside the domain and how this impacts broader concerns of fairness and trust. Future research must connect specialized contributions back to a holistic socio-technical perspective on generative AI systems to ensure that disciplinary specializations remain impactful and aligned with the complexity of generative AI systems. This connection involves actively considering diverse stakeholder perceptions, the behavioral implications of interacting with generative AI, and the broader institutional, cultural, and legal contexts in which organizations deploy such systems. Only through such integration can an accountability framework for generative AI systems that is technically rigorous, domain-appropriate, socially embedded, and globally relevant be created. In light of these considerations, we propose our final research questions for our research agenda:

> ***RA3.1***: *How can accountability mechanisms for generative AI systems be designed to reflect both technical affordances and social expectations across diverse application contexts?*

> ***RA3.2***: *To what extent do disciplinary specializations hinder or promote the development of holistic and socio-technical perspectives on accountability for generative AI systems?*

Table 3 Research Agenda for Accountability for Generative AI Systems

Research Proposition		Derived Research Question for the Research Agenda for Accountability for Generative AI Systems	
No	**Proposition**	**No**	**Research Question**
1	Research disciplines should strive to balance external integration with internal consolidation by critically engaging with their own prior work to develop coherent and distinctive perspectives on accountability for AI systems.	RA1.1	How do the accountability challenges of generative AI systems reshape foundational assumptions of accountability for AI systems within individual disciplines?
		RA1.2	How can disciplinary perspectives be combined to create a coherent and cumulative understanding of accountability for generative AI systems?
		RA1.3	How can disciplines engage more effectively in reciprocal interdisciplinary discourse on accountability for generative AI systems once a strong disciplinary foundation is established?
2	Future research should consider the different conceptualizations of accountability for AI systems (i.e., as a governance mechanism or a virtue) when transferring knowledge across disciplinary boundaries.	RA2.1	How do different conceptualizations of accountability for AI systems, such as governance mechanisms or virtues, shape the development, deployment, and regulation of generative AI systems?
		RA2.2	How can governance- and virtue-based conceptualizations of accountability for AI systems be integrated to create comprehensive accountability frameworks for managing generative AI systems?
		RA2.3	Which accountability mechanisms are most effective in shaping accountability as a governance mechanism or as a virtue to govern generative AI systems?
3	Future research should integrate domain-specific and methodological specializations into a broader socio-technical perspective to ensure a more comprehensive understanding of accountability for AI systems.	RA3.1	How can accountability mechanisms for generative AI systems be designed to reflect both technical affordances and social expectations across diverse application contexts?
		RA3.2	To what extent do disciplinary specializations hinder or promote the development of holistic and socio-technical perspectives on accountability for generative AI systems?
		RA3.3	How do stakeholder perceptions of accountability for generative AI systems differ across domains, and how should these differences inform the design and governance of generative AI systems?

> ***RA3.3****: How do stakeholder perceptions of accountability for generative AI systems differ across domains, and how should these differences inform the design and governance of generative AI systems?*

Addressing RA3.1, researchers might explore how technical applications and tools, such as explanation interfaces (e.g., Ribeiro et al., 2016) or AI audits (e.g., Raji et al., 2020) are interpreted differently by generative AI systems' stakeholders, such as healthcare professionals or regulators. Methods from different disciplines, such as IS and psychology, can help uncover how these interfaces or AI audits shape stakeholders' perceptions of accountability for generative AI systems. Addressing RA3.2, researchers can examine how disciplinary specializations, such as legal perspectives emphasizing liability (e.g., Novelli et al., 2024) versus human-centered design approaches prioritizing user empowerment (e.g., Schmager et al., 2025), may either constrain or support the development of more holistic, socio-technical perspectives on accountability for generative AI systems. Finally, addressing RA3.3, researchers might be aware that different users, such as healthcare professionals, regulators, or developers, might interpret accountability through their professional norms and values. These differing interpretations might influence how they perceive and handle accountability mechanisms for generative AI systems. In this context, researchers could conduct domain-specific case studies or cross-sectoral surveys to uncover these differences in interpretations. Understanding and incorporating stakeholder perspectives can help tailor accountability frameworks for generative AI systems to be contextually appropriate.

5.3.3 Summary of the Research Agenda for Accountability for Generative AI Systems

Table 3 summarizes the research agenda derived from our analysis, which aims to guide future investigations into accountability for generative AI systems. This agenda emphasizes the importance of encouraging disciplines to strengthen and consolidate their internal theoretical foundations before engaging in interdisciplinary discourse (RA1) and highlights the need to consider varying conceptualizations of accountability for AI systems as either a governance mechanism or a virtue when designing accountability frameworks (RA2), and advocates for reconnecting domain-specific and technical specializations to a broader socio-technical understanding of accountability for generative AI systems (RA3).

5.4 Implications for Practice

Our study provides practical implications for policymakers and organizations facing accountability for AI or generative AI systems. First, we give a synthesized and aggregated view of existing research on accountability for AI systems.

As a result, we offer policymakers a comprehensive overview and starting point to better access existing research on accountability for AI systems. This synthesized and aggregated view is crucial for policymakers to make informed decisions when shaping regulations, as it helps them to understand the implications of accountability for AI systems. Therefore, policymakers can better integrate accountability for AI systems into future planned regulations, such as the Algorithmic Accountability Act in the United States and further improvements on the already introduced AI Act of the European Union (117th Congress USA, 2022; Council of the European Union, 2024).

Second, we highlight the necessity for organizations to recognize that research on accountability for AI systems understands it as a governance mechanism or a virtue. This understanding is essential for organizations since they can use accountability as a governance mechanism or a virtue within their AI development projects. When using accountability as a governance mechanism, organizations can implement clear guidelines and rules to enforce explanation and justification by developers when their AI systems cause harm. Moreover, when using accountability as a virtue, organizations must highlight accountability for AI systems in front of their developers during AI development projects. This approach enables organizations to motivate developers who design, implement, and deploy AI systems to consider their accountability beforehand and to reduce potential harm by preventing the potential negative impacts of their AI systems. As a result, drawing on these two perspectives enables organizations to benefit from accountability across the whole life cycle of their AI systems.

5.5 Limitations

Our study has limitations that offer valuable starting points to enhance our observation of research on accountability for AI systems. First, since we primarily relied on the Web of Science database to provide the necessary metadata for our analysis, our dataset depended on only one scientific database. Future research might integrate other scientific databases like Scopus to broaden research coverage. Future research might also consider other databases, such as social media posts and newspaper articles, to understand how the scientific understanding of accountability for AI systems and generative AI systems translates to the public.

Second, we divided the extracted publications based on the inherent Web of Science Category *Computer Science Information Systems*, relying on an independent non-IS and IS research classification. Future research might use other classifications, such as the German VHB ranking, to divide the extracted publications into non-IS and IS research. The use of such rankings would put more emphasis on geographical regions or specific research communities (e.g., German-speaking countries for the VHB ranking). This emphasis on geographical regions would allow research to understand the nuances of accountability for AI systems and generative AI systems within geographic areas and specific communities, and would contribute to the overall robustness of our study.

Third, using our two datasets, we created *Word2Vec* word embeddings for similarity analysis. Since the number of words in our two datasets is relatively small, the word embeddings may have limited robustness. Future research could enhance robustness by analyzing full publications. Additionally, employing alternative classical analysis methods, such as Levenshtein distance, could provide valuable comparative data, further supporting the validity of our findings regarding the understanding of accountability for AI systems and generative AI systems.

6 Conclusion

Given the unprecedented growth in AI systems and, in particular, generative AI systems across various applications within organizations and society, ensuring accountability to guide their safe and responsible usage is crucial. However, much research on accountability for AI systems exists without synthesizing and aggregating their findings, making it challenging to access this research. Against this backdrop, this paper synthesizes and aggregates non-IS and IS research findings to observe what we can learn from existing research on accountability for AI systems. In doing so, our paper provides researchers better access to research on accountability for AI systems. Additionally, it highlights pathways for future research to gain a more profound understanding of accountability for AI systems to promote future designs and implementations of AI and generative AI systems guided by accountability.

Acknowledgements Not applicable.

Authors' Contributions Conceptualization: Sebastian Clemens Bartsch; Methodology: Sebastian Clemens Bartsch, Long Hoang Nguyen; Formal analysis and investigation: Sebastian Clemens Bartsch, Long Hoang Nguyen; Writing—original draft preparation: Sebastian Clemens Bartsch, Long Hoang Nguyen; Writing—review and editing: Jan-Hendrik Schmidt, Guangyu Du; Funding acquisition: Alexander Benlian, Ali Sunyaev; Resources: Sebastian Clemens Bartsch; Supervision: Jan-Hendrik Schmidt, Martin Adam, Alexander Benlian, Ali Sunyaev.

Funding Open Access funding enabled and organized by Projekt DEAL. We gratefully acknowledge funding for this work by the German Research Foundation (DFG) through the granted project "Accountable Artificial Intelligence-based Systems" (project number 471168026).

Data Availability We used public data that anyone can access for our analyses. We have included a list of the publications we have considered in the online appendix.

Declarations

Ethics Approval Not applicable.

Competing interests No competing financial and non-financial interests exist.

Others During the preparation of this work, the authors used ChatGPT-4o (chatgpt.com) to improve language and readability. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

117th Congress USA. (2022). *H.R.6580 - Algorithmic Accountability Act of 2022*. Retrieved 04/10/2023 from https://www.congress.gov/bill/117th-congress/house-bill/6580/text

Adam, M. (2022). Accountability-based user interface design artifacts and their implications for user acceptance of AI-Enabled Services. 2022 European Conference on Information Systems, Timișoara, Romania.

Adam, M., Roethke, K., & Benlian, A. (2023). Human versus automated sales agents: How and why customer responses shift across sales stages. *Information Systems Research, 34*(3), 1148–1168.

Ågerfalk, P., Conboy, K., Crowston, K., Eriksson Lundström, J., Jarvenpaa, S., Mikalef, P., & Ram, S. (2021). Artificial intelligence in information systems: State of the art and research roadmap. *Communications of the Association for Information Systems, 50*(1), 420–438.

Aleksovska, M., Schillemans, T., & Grimmelikhuijsen, S. (2019). Lessons from five decades of experimental and behavioral research on accountability: A systematic literature review. *Journal of Behavioral Public Administration, 2*(2), 1–18.

Appelbaum, S. H. (1997). Socio-technical systems theory: An intervention strategy for organizational development. *Management Decision,35*(6), 452–463.

Armstrong, E. (2005). Integrity, Transparency and Accountability in Public Administration: Recent Trends, Regional and International Developments and Emerging Issues. United Nations, *Department of Economic and Social Affairs, 1*(10), 1–14.

Bartsch, S. C., Behn, O., Benlian, A., Brownsword, R., Bücker, S., Düwell, M., Formánek, N., Jungtäubl, M., Leyer, M., Richter, A., Schmidt, J.-H., & Will-Zocholl, M. (2025). Governance of AI Systems in High-Risk Contexts. *Business & Information Systems Engineering.*

Bartsch, S. C., Milani, V., Adam, M., & Benlian, A. (2024). *Algorithmic Accountability: What* Does *It Mean for AI Developers and How Does It Affect AI Development Projects*. Hawaii International Conference on System Sciences, O'ahu, USA.

Bartsch, S. C., & Schmidt, J.-H. (2023). *How AI Developers' Perceived Accountability Shapes Their AI Design Decisions*. International Conference on Information Systems, Hyderabad, India.

Batey, M. V., & Lewis, F. M. (1982). Clarifying autonomy and accountability in nursing service: Part 1. *Journal of Nursing Administration, 12*(9), 13–18.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly, 45*(3), 1433–1450.

Berger, B., Adam, M., Rühr, A., & Benlian, A. (2021). Watch me improve - Algorithm aversion and demonstrating the ability to learn. *Business & Information Systems Engineering, 63*(1), 55–68.

Beu, D., & Buckley, M. R. (2001). The hypothesized relationship between accountability and ethical behavior. *Journal of Business Ethics, 34*, 57–73.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal, 13*, 447–468.

Bovens, M. (2010). Two concepts of accountability: Accountability as a virtue and as a mechanism. *West European Politics, 33*(5), 946–967.

Chan, J., Peko, G., Sundaram, D., & Sadovykh, V. (2025). Application and adoption of emerging technologies: From AI to metaverse. *Information Systems Frontiers, 27*, 1–4.

Checkland, K., Marshall, M., & Harrison, S. (2004). Re-thinking accountability: Trust versus confidence in medical practice. *BMJ Quality & Safety, 13*(2), 130–135.

Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. *Washington Law Review, 89*, 1–33.

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management, 60*, Article 102383.

Cooper, A. F., Moss, E., Laufer, B., & Nissenbaum, H. (2022). *Accountability in an Algorithmic Society: Relationality, Responsibility, and Robustness in Machine Learning* 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea.

Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, 2024 Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, 1–272 (2024). https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf

Dastin, J. (2022). Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women. In *Ethics of Data and Analytics* (pp. 296–299). Auerbach Publications.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285–296.

Du, G., Lins, S., Blohm, I., & Sunyaev, A. (2024). *My Fault, Not AI's Fault. Self-Serving Bias Impacts Employees' Attribution of AI* Accountability International Conference of Information Systems, Bangkok, Thailand.

Dubnick, M. (2005). Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review, 28*(3), 376–417.

Durante, M. (2013). How to cross boundaries in the information society: Vulnerability, responsiveness, and accountability. *ACM SIGCAS Computers and Society, 43*(1), 9–21.

Emanuel, E. J., & Emanuel, L. L. (1996). What is accountability in health care? *Annals of Internal Medicine, 124*(2), 229–239.

Ferrara, E. (2024). Genai against humanity: Nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science, 7*(1), 549–569.

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

Fosso Wamba, S., & Queiroz, M. M. (2023). Responsible artificial intelligence as a secret ingredient for digital health: Bibliometric analysis, insights, and research directions. *Information Systems Frontiers, 25*(6), 2123–2138.

French, A. M., & Shim, J. (2024). From Artificial Intelligence to Augmented Intelligence: A Shift in Perspective, Application, and Conceptualization of AI. *Information Systems Frontiers*, 1–22.

Gal, U., Hansen, S., & Lee, A. S. (2022). Research perspectives: Toward theoretical rigor in ethical analysis: The case of algorithmic decision-making systems. *Journal of the Association for Information Systems, 23*(6), 1634–1661.

Giesen, I., & Kristen, F. G. (2014). Liability, responsibility and accountability: Crossing borders. *Utrecht Law Review, 10*(3), 1–13.

Gupta, M., Parra, C. M., & Dennehy, D. (2022). Questioning racial and gender bias in AI-based recommendations: Do espoused national cultural values matter? *Information Systems Frontiers, 24*(5), 1465–1481.

Hall, A. T., Frink, D. D., & Buckley, M. R. (2017). An accountability account: A review and synthesis of the theoretical and empirical research on felt accountability. *Journal of Organizational Behavior, 38*(2), 204–224.

Hammerschmidt, T., Hafner, A., Stolz, K., Passlack, N., Posegga, O., & Gerholz, K.-H. (2025). A Review of How Different Views on Ethics Shape Perceptions of Morality and Responsibility within AI Transformation. *Information Systems Frontiers*, 1–87.

Heinisuo, E., Kuoppakangas, P., & Stenvall, J. (2025). Navigating AI Implementation in Local Government: Addressing Dilemmas by Fostering Mutuality and Meaningfulness. *Information Systems Frontiers*, 1–22.

High-Level Expert Group on AI. (2019). *Ethics Guidelines for Trustworthy AI*. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Hill, K. (2023). Eight Months Pregnant and Arrested after False Facial Recognition Match. *The New* York *Times*. https://www.nytimes.com/2023/08/06/business/facial-recognition-false-arrest.html

Horneber, D., & Laumer, S. (2023). Algorithmic accountability. *Business & Information Systems Engineering*. https://doi.org/10.1007/s12599-023-00817-8

Hughes, L., Malik, T., Dettmer, S., Al-Busaidi, A. S., & Dwivedi, Y. K. (2025). Reimagining Higher Education: Navigating the Challenges of Generative AI Adoption. *Information Systems Frontiers*, 1–23.

IBM (2022). *Accountability*. Retrieved 01/07/2023 from https://www.ibm.com/design/ai/ethics/accountability/

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*, 389–399.

Jussupow, E., Spohrer, K., & Heinzl, A. (2022). Radiologists' usage of diagnostic AI systems: The role of diagnostic self-efficacy for sensemaking from confirmation and disconfirmation. *Business & Information Systems Engineering, 64*(3), 293–309.

Kahdan, M., Hartwich, N. J., & Salge, O. (2022). *Ethical AI Research Untangled: Mapping Interdisciplinary Perspectives for Information Systems Research* International Conference on Information Systems (ICIS), Copenhagen, Denmark.

Katz, U., Cohen, E., Shachar, E., Somer, J., Fink, A., Morse, E., Shreiber, B., & Wolf, I. (2024). GPT versus resident physicians—a benchmark based on official board scores. *NEJM AI, 1*(5), 1–8.

Kenthapadi, K., Lakkaraju, H., & Rajani, N. (2023). *Generative AI Meets Responsible AI:* Practical *Challenges and Opportunities* ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Long Beach, USA.

Kim, B., Park, J., & Suh, J. (2020). Transparency and accountability in AI decision support: Explaining and visualizing convolutional neural networks for text information. *Decision Support Systems,134*(113302), 1–11.

Latino, M. E., Menegoli, M., & Corallo, A. (2022). Agriculture digitalization: A global examination based on bibliometric analysis. *IEEE Transactions on Engineering Management, 71*, 1330–1345.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*(2), 255–275.

Li, Y., & Goel, S. (2024). Making it possible for the auditing of AI: A systematic review of AI audits and AI auditability. *Information Systems Frontiers, 27*, 1121–1151. https://doi.org/10.1007/s10796-024-10508-8

Li, Y., Gu, T., Yang, C., Li, M., Wang, C., Yao, L., Gu, W., & Sun, D. (2025). AI-assisted hypothesis generation to address challenges in cardiotoxicity research: Simulation study using ChatGPT with GPT-4o. *Journal of Medical Internet Research, 27*(e66161), 1–19.

Marabelli, M., & Newell, S. (2023). Responsibly strategizing with the metaverse: Business implications and DEI opportunities and challenges. *Journal of Strategic Information Systems, 32*(2), Article 101774.

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics, 160*, 835–850.

Microsoft Corporation. (2022). *Microsoft Responsible AI Standard, V2*. Microsoft Corporation. Retrieved 01/07/2023 from https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE5cmFl

Mikalef, P., Benlian, A., Conboy, K., & Tarafdar, M. (2025). Responsible AI starts with the artifact: Challenging the concept of responsible AI in IS research. *European Journal of Information Systems, 34*(3), 407–414.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. https://arxiv.org/abs/1301.3781, 1–12.

Moss, E. (2023). Trust is not enough: Accuracy, error, randomness, and accountability in an algorithmic society. *Communications of the ACM, 66*(6), 42–44.

Mulgan, R. (2000). 'Accountability': An Ever-Expanding Concept? *Public Administration, 78*(3), 555–573.

Naja, I., Markovic, M., Edwards, P., Pang, W., Cottrill, C., & Williams, R. (2022). Using knowledge graphs to unlock practical collection, integration, and audit of AI accountability information. *IEEE Access, 10*, 74383–74411.

Nguyen, L. H., Lins, S., Renner, M., & Sunyaev, A. (2024). *Unraveling the Nuances of AI Accountability: A Synthesis of Dimensions across Disciplines* European Conference on Information Systems, Paphos, Cyprus.

Nguyen, T., & Elbanna, A. (2025). Understanding Human-AI Augmentation in the Workplace: A Review and a Future Research Agenda. *Information Systems Frontiers*, 1–21.

Niu, S., Ma, J., Yin, Q., Wang, Z., Bai, L., & Yang, X. (2024). Modelling patient longitudinal data for clinical decision support: A case study on emerging AI healthcare technologies. *Information Systems Frontiers, 27*, 409–427.

Novelli, C., Casolari, F., Hacker, P., Spedicato, G., & Floridi, L. (2024). Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity. *Computer Law & Security Review, 55*, Article 106066.

Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: What it is and how it works. *AI and Society, 1*, 1–12.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453.

Ogie, R. I., O'Brien, S., & Federici, F. M. (2022). Towards using agent-based modelling for collaborative translation of crisis information: A systematic literature review to identify the underlying attributes, behaviours, interactions, and environment of agents. *International Journal of Disaster Risk Reduction, 68*, Article Article 102717.

Painter-Morland, M. (2006). Redefining accountability as relational responsiveness. *Journal of Business Ethics, 66*, 89–98.

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems, 34*(2), Article 101885.

Pichai, S. (2018). *AI at Google: Our Principles*. Google. Retrieved 01/07/2023 from https://blog.google/technology/ai/ai-principles/

Pinski, M., & Benlian, A. (2024). AI literacy for users–A comprehensive review and future research directions of learning methods, components, and effects. *Computers in Human Behavior: Artificial Humans,2*(1), 100062.

Rai, A., Constantinides, P., & Sarker, S. (2019). Next generation digital platforms: Toward human-AI hybrids. *MIS Quarterly, 43*(1), iii–ix.

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?" Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 1135–1144.

Sahoo, S., Kumar, A., Mishra, R., & Tripathi, P. (2022). Strengthening supply chain visibility with blockchain: A Prisma-based review. *IEEE Transactions on Engineering Management, 71*, 1787–1803.

Samtani, S., Zhao, Z., & Krishnan, R. (2023). Secure knowledge management and cybersecurity in the era of artificial intelligence. *Information Systems Frontiers, 25*(2), 425–429.

Sanderson, C., Douglas, D., Lu, Q., Schleiger, E., Whittle, J., Lacey, J., Newnham, G., Hajkowicz, S., Robinson, C., & Hansen, D. (2023). AI ethics principles in practice: Perspectives of designers and developers. *IEEE Transactions on Technology and Society, 4*(2), 171–187.

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance. *MIS Quarterly, 43*(3), 695–720.

Schmager, S., Pappas, I. O., & Vassilakopoulou, P. (2025). Understanding Human-Centred AI: A Review of Its Defining Elements and a Research Agenda. *Behaviour & Information Technology*, 1–40.

Schmidt, J.-H., Bartsch, S. C., Adam, M., & Benlian, A. (2025). Elevating developers' accountability awareness in AI systems development: The role of process and outcome accountability arguments. *Business & Information Systems Engineering, 67*, 109–135.

Schöbel, S., Schmitt, A., Benner, D., Saqr, M., Janson, A., & Leimeister, J. M. (2023). Charting the evolution and future of conversational agents: A research agenda along five waves and new frontiers. *Information Systems Frontiers, 26*(2), 729–754.

Söllner, M., Arnold, T., Benlian, A., Bretschneider, U., Knight, C., Ohly, S., Rudkowski, L., Schreiber, G., & Wendt, D. (2025). ChatGPT and beyond: Exploring the responsible use of generative AI in the workplace: An interdisciplinary perspective. *Business & Information Systems Engineering, 67*, 289–303.

Stafford, A., & Stapleton, P. (2022). The impact of hybridity on PPP governance and related accountability mechanisms: The case of UK education PPPs. *Accounting, Auditing & Accountability Journal,35*(3), 950–980.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science,19*(3), 387–420.

Storey, V. C., Yue, W. T., Zhao, J. L., & Lukyanenko, R. (2025). Generative Artificial Intelligence: Evolving Technology, Growing Societal Impact, and Opportunities for Information Systems Research. *Information Systems Frontiers*, 1–22.

Sunyaev, A., Benlian, A., Pfeiffer, J., Jussupow, E., Thiebes, S., Maedche, A., & Gawlitza, J. (2025). High-Risk Artificial Intelligence. *Business & Information Systems Engineering*, 1–14.

Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. *Electronic Markets, 31*, 447–464.

van Eck, N. J., & Waltman, L. (2018). Vosviewer Manual V1.6.8. In: Universiteit Leiden.

Vössing, M., Kühl, N., Lind, M., & Satzger, G. (2022). Designing transparency for effective human-AI collaboration. *Information Systems Frontiers, 24*(3), 877–895.

Wan, Y., Pu, G., Sun, J., Garimella, A., Chang, K.-W., & Peng, N. (2023). "Kelly Is a Warm Person, Joseph Is a Role Model": Gender Biases in LLM-Generated Reference Letters. *arXiv Preprint*, 1–19.

Wieringa, M. (2020). *What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability* Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency, Barcelona, Spain.

Wieringa, M. (2023). "Hey Syri, Tell Me About Algorithmic Accountability": Lessons from a Landmark Case. *Data & Policy, 5*, 1–24.

Xia, B., Lu, Q., Zhu, L., Lee, S. U., Liu, Y., & Xing, Z. (2024). *Towards a Responsible AI Metrics Catalogue: A Collection of Metrics for AI Accountability*. IEEE/ACM International Conference on AI Engineering-Software Engineering for AI, Lisbon, Portugal.

Zemčík, T. (2021). Failure of chatbot Tay was evil, ugliness and uselessness in its nature or do we judge it through cognitive shortcuts and biases? *AI and Society, 36*(1), 361–367.

Zhong, C., & Yayla, A. (2025). Cognitive Impacts of Explainable AI in Cybersecurity Incident Response: Challenges and Propositions. *Information Systems Frontiers*.

Sebastian Clemens Bartsch is a Ph.D. candidate at the chair of Information Systems & E-Services at the Technical University of Darmstadt, Germany. He holds a Master's degree in Business Information Systems from the Technical University of Darmstadt. His research interests include accountability in the context of AI-based information systems and explainable AI. His work has been published at international Information Systems conferences such as the International Conference on Information Systems, Artificial Intelligence, Ethics & Society, and the Hawaii International Conference on Systems Science.

Long Hoang Nguyen is a research associate at the Karlsruhe Institute of Technology and a Ph.D. candidate at the Technical University of Munich, Germany. He holds a Master's degree in Information Systems from the University of Potsdam. His research primarily concerns the socio-technical means to facilitate accountability for AI-based information systems. His work has appeared at leading international Information Systems conferences, such as the European Conference on Information Systems.

Jan-Hendrik Schmidt is a research associate at the Chair of Information Systems & E-Services at the Technical University of Darmstadt, Germany, where he also earned his Ph.D. in Information Systems Management. His research primarily focuses on accountable AI-based information systems and methodological advances within the Information Systems research discipline. He employs a range of quantitative and qualitative research methods, and his work has appeared in the journal Business & Information Systems Engineering as well as in the proceedings of major Information Systems conferences (e.g., International Conference on Information Systems, European Conference on Information Systems, Hawaii International Conference on Systems Science).

Guangyu Du is a research associate at the Karlsruhe Institute of Technology and a Ph.D. candidate at the Technical University of Munich, Germany. She holds a Master's degree from New York University. Her research focuses on accountable AI-based information systems, particularly users' perceptions and behavior. Her work has appeared at leading international Information Systems conferences, such as the International Conference on Information Systems.

Martin Adam is a chaired professor of Information Systems, especially Smart Services, at the University of Goettingen, Germany. His research interests include user-AI collaborations as well as the digital transformation of work and people. Martin Adam's work has been published in journals such as Information Systems Research, Journal of the Association for Information Systems, Information Systems Journal, European Journal of Information Systems, and others. He serves as a Senior Editor for Electronic Markets and as an associate editor for Information Systems Journal and Business & Information Systems Engineering.

Alexander Benlian is a chaired professor of Information Systems & E-Services at the Technical University of Darmstadt, Germany. His research focuses on algorithmic management, accountable AI, IT entrepreneurship, and human-computer interaction. His work has been published in journals such as MIS Quarterly, Information Systems Research, Journal of Management Information Systems, Entrepreneurship Theory & Practice, Journal of Business Venturing, and Journal of the Association for Information Systems. Alex serves as senior editor at the European Journal of Information Systems, associate editor at Information Systems Research, and as department editor at Business & Information Systems Engineering. He also serves on the editorial boards of MIS Quarterly Executive and Electronic Markets.

Ali Sunyaev is a chaired professor of Information Systems / Computer Science at the Technical University of Munich, Germany. Before joining the Technical University of Munich, he was a chaired professor at the Karlsruhe Institute of Technology, the University of Kassel, and the University of Cologne. His research work accounts for the multifaceted use contexts of digital technologies, with research on human behavior affecting IT applications and vice versa. His research appeared in journals, including Information Systems Research, Journal of Management Information Systems, Journal of Information Technology, Journal of the Association for Information Systems, IEEE Transactions on Cloud Computing, Communications of the ACM, and others.