

Next-generation IS research methods – towards a better understanding of complex and dynamic phenomena ... and generative AI as the elephant in the room

Journal of Information Technology
2025, Vol. 40(2) 102–121
© Association for Information
Technology Trust 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/02683962251340699

journals.sagepub.com/jinf



Ivo Blohm¹ , Shaila Miranda², Shuk Ying Ho³  and Jan Marco Leimeister^{1,4}

Abstract

Emerging phenomena like the increasing volume and variety of available data, machine learning techniques, and, most recently, Generative AI are reshaping our research practices, affording new research methods for testing and developing theory. In this editorial, we discuss our two major observations from running the ‘Next-generation IS Research Methods’ Special Issue: (1) the need for research methods that enhance our understanding of complex and dynamic phenomena and (2) Generative AI as (potential) productivity enhancer. We compile these observations into an organizing framework and discuss possibilities for applying Generative AI in the fields of qualitative, quantitative, and engaged research. We highlight challenges that might occur when applying Generative AI in research and shed light on the changing role of researchers in such settings of human–AI collaboration.

Keywords

Qualitative research, quantitative research, engaged research, research methods, generative artificial intelligence, computational theory construction

Introduction

Information systems (IS) researchers have applied a range of quantitative, qualitative, and engaged methods to study information technology (IT)-related phenomena. Quantitative IS research has typically followed a positivist approach of hypothesis testing, with ‘first generation’ regression models distinguished from post-2000 ‘second generation’ structural equation models (Gerow et al., 2011); data used in such research stems from sources such as surveys, experiments, and panel studies. The primary objective of this research stream typically is theory testing. Qualitative IS research has ranged from positivist to interpretivist, incorporating deductive to inductive/abductive reasoning (Sarker et al., 2018). Sarker et al. (2018) detailed five major qualitative methods genres: positivist case studies, grounded theory methodologies, exploratory case studies, interpretive case studies/ethnographies, and hermeneutics. The primary objective of this research stream is theory construction. A third type of IS research seeks new knowledge with ‘engaged’ methods like design science research (DSR) or action research. The objective of this

stream that differentiates it from the first two is the focus on solving important individual, organisational, or societal challenges and/or extending the boundaries of human and organisational capabilities by solving real-world problems, striving for usefulness, and creating new and innovate artefacts (Baskerville and Wood-Harper, 1996; Tuunanen et al., 2024).

In recent years, the scope of and possibilities for IS research have expanded. ‘Digital’ IT have become mainstream at the individual, organisational, and societal level

¹University of St.Gallen, Institute for Information Systems and Digital Business, Switzerland

²University of Arkansas, Walton College of Business, USA

³Australian National University, College of Business and Economics, Australia

⁴Kassel University, Chair for Information Systems, Germany

Corresponding author:

Ivo Blohm, Institute of Information Systems and Digital Business, University of St.Gallen, Müller-Friedberg-Strasse 8, St.Gallen 9000, Switzerland.

Email: ivo.blohm@unisg.ch

(Burton-Jones et al., 2021). Digital IT is the key to innovation in domains from medicine to education, and from psychology to the arts. Hence, these domains have become of interest and relevance to IS researchers, yet also interdisciplinary in nature. These innovations are increasing the volume and variety of data available to researchers but have necessitated shifts in our research practices (Johnson et al., 2019). Further, big data, machine learning (ML), artificial intelligence (AI), and most recently Generative AI (GAI) technologies provide a plethora of powerful approaches for collecting and analysing or even creating data at a scale not possible before. This requires us to adapt our research methods and tools to make this data actionable and to generate novel knowledge.

The ways that novel technology capacities afford new possibilities is quite apparent in quantitative IS research. In contrast to ‘first generation’ regression-based approaches and ‘second generation’ structural equation modelling, which are based on hypothesise-then-test, a new generation of research methods including predictive analytics (Shmueli and Koppius, 2011), data mining (Smith, 2020), ML (Shrestha et al., 2020), or explainable AI (XAI; Gunning et al., 2019) confer substantial new opportunities and challenges. For instance, these methods potentially allow for more fine-grained measurements and analyses to extend our knowledge of existing phenomena and permit access to novel phenomena that were previously out of reach (George et al., 2016). These technological advancements enable us to progress our research toolkit and inform new ways of generating knowledge and theorising (Burton-Jones et al., 2021; Shrestha et al., 2020).

Qualitative-interpretative research is also applying new methods suitable to the wider, interdisciplinary nature of IS research. Since Sarker et al. (2018)’s review of ‘first generation’ qualitative methods, qualitative researchers have also leveraged emerging computational techniques to take advantage of the proliferating trace data sets through the first two decades of the 21st century (Lindberg, 2020). These approaches, referred to as computational theory construction (CTC), entail application of computational techniques for pattern surfacing toward theory construction (Berente et al., 2019; Miranda et al., 2022a).

Similarly, today’s pace of innovation also warrants novel approaches to engaged scholarship and design science. For instance, researchers have been highlighting that design science research involving ML/AI invokes a specific set of complexities researchers have to deal with when designing AI-enabled IS (Abbasi et al., 2024). While ML/AI have an increasing impact on our research methods toolbox and offer new research opportunities, their wide-spread application also poses ethical and societal questions for researchers. This is particularly true for Large Language Models (LLM) or, more generally, GAI, the proliferation of which have started debates about scholarly human-AI

collaboration and the (potentially) changing role of researchers in the scientific process of knowledge discovery (e.g. Davison et al., 2024; Grimes et al., 2023; Ngwenyama and Rowe, 2024; Schlagwein and Willcocks, 2023; Stokel-Walker and Van Noorden, 2023; Van Dis et al., 2023).

This special issue sought to elicit a new generation of IS research methods that account for the shift from IT to digital (e.g. Piccoli et al., 2022) and the interdisciplinary nature and wider scope of phenomena involving information and digital technologies and the novel capacities afforded by new technologies/techniques (such as ML/AI). Our special issue call for papers preceded the explosion of GAI and was therefore too early to capture contributions exploring GAI support for conducting research. Nonetheless, we aim to connect the general trends that the proposed next-generation IS research methods in this special issue unravel to the current scholarly debate on GAI. Consequently, this editorial not only presents complements to today’s methodological toolbox of IS researchers, but also discusses the potential affordances that GAI might bring for the future of quantitative, qualitative, and engaged research in the IS domain. We first provide a synopsis of the articles in the special issue, synthesizing them into an organising framework for next-generation IS research methods. Next, we discuss applications of GAI in the three dominant research paradigms, how we might overcome known challenges of GAI-enabled research such as hallucination, as well as the (potentially) changing role of researchers in human–AI collaboration in the scientific knowledge creation process.

Synopsis of the articles in the special issue

In this section, we briefly summarise the next-generation IS research methods proposed in articles that we accepted for the special issue. We highlight the methodological areas that the author teams strive to advance and their approach to improve existing methods.

The first article by Mechthild Pieper, Monica Fallon, and Armin Heinzl on ‘Micro-Randomized Trials in Information Systems Research: An Experimental Method for Advancing Knowledge about our Dynamic and Digitalized World’ critiques static experimental designs for their inability to capture the dynamic and complex phenomena that characterise the focus of current IS scholarship. These static methods fail to address the evolving interactions between variables and their impacts over time. Static designs often ignore context-specific factors that alter outcomes, leading to incomplete conclusions. The lack of temporal granularity restricts understanding of how interventions work over time, how they interact with changing contexts, and how they can be adapted to improve effectiveness continuously. To address these limitations, the authors propose using micro-randomized trials (MRTs) as an advanced

experimental design for IS research. MRTs evaluate just-in-time adaptive interventions, allowing dynamic assessment of treatment effects over time. This methodology enables data collection and analysis at multiple time points, providing richer insights into contextual factors influencing intervention effectiveness. MRTs help identify causal relationships between variables and their covariates, enhancing the precision of IS research through continuous feedback and adaptation. Pieper et al. illustrate MRTs through a case study involving a mobile health intervention aimed at encouraging physical activity. Their demonstration includes the design, implementation, and evaluation phases of the MRT. Specifically, the study dynamically assesses the intervention's impact on user behaviour over time, highlighting the iterative nature of MRTs. Real-time data collection and analysis lead to continuous improvement and adaptation of the intervention, providing deeper insights into factors driving engagement and the effectiveness of adaptive interventions in a real-world setting.

Shohil Kishore, David Sundaram, and Michael Myers contributed the second article of the special issue on 'Temporal Dynamics Framework and Methodology for Social Media Research'. The paper criticises traditional qualitative and quantitative methods in IS research for their inability to effectively integrate temporality and processual dynamics in the analysis of textual trace data in the social media context. Though theoreticians have issued many calls for careful attention to temporal dynamics in IS research (e.g. Saunders et al., 2004), traditional IS research methodologies often fail to capture the temporal and processual dynamics inherent in trace data in social media and beyond. These methods typically focus on cross-sectional or static analyses, insufficient for understanding how social media use and effects change over time. This oversight leads to fragmented insights that do not fully reflect continuous and dynamic processes driving social media behaviour. The lack of temporal integration restricts comprehensive process theory development from social media trace data and beyond. Kishore, Sundaram, and Myers propose the 'Temporal Dynamics Framework and Methodology' (TDFM) which emphasises integrating temporality in data collection and analysis, enabling process theory construction. This framework guides iterative data collection, analysis, and temporal contextualization, addressing unique challenges and opportunities in social media research. TDFM builds on computationally intensive methods, combining manual and automated approaches to construct theory from dynamic, time-stamped social media data. TDFM is demonstrated through a study of mental health discourse on Twitter during COVID-19. The framework guides data collection and analysis, revealing dynamic shifts in social media use across pandemic phases. By incorporating temporal dynamics, the study provides deeper insights into psychological needs motivating social media use and how these needs evolve.

The case study highlights TDFM's practical application, showcasing its ability to enhance process theory construction by integrating temporality in research using qualitative trace data.

The third article by Dominik Stoffels, Stefan Faltermaier, Kim Strunk, and Marina Fiedler on 'Guiding Computationally Intensive Theory Development with Explainable Artificial Intelligence: The Case of SHAP' points out existing post-hoc XAI methods, particularly SHAP (SHapley Additive exPlanations), for limitations in mapping ML patterns to support CTC in IS research. Existing XAI methods, including SHAP, are criticised for oversimplified and sometimes inaccurate representations of patterns learned by ML algorithms. These methods often fail to capture the complexity of interaction effects and high-dimensional datasets, leading to misleading interpretations. Such inaccuracies distort understanding phenomena and compromise theory development validity. Specific shortcomings include SHAP's misrepresentation of variable influence in the presence of interaction effects and oversimplification in visualisations. Stoffels et al. propose a Monte Carlo simulation to evaluate SHAP's performance in uncovering meaningful patterns for CTC. Additionally, an empirical case study applies their application guidelines to real-world data, identifying conditions for valuable insights. The approach provides guidelines for using SHAP effectively in CTC. The approach is demonstrated through applications of SHAP to synthetic and real-world datasets. In synthetic analyses, SHAP identifies underlying patterns, compared against known ground truths for accuracy. The empirical case study analyses patterns in airline customer satisfaction, discussing implications for theory development. The demonstration showcases SHAP's practical benefits and limitations, providing concrete examples of effective XAI tool use in CTC. Findings highlight conditions where SHAP can uncover meaningful patterns, enhancing IS theories' robustness.

The fourth article by Annamina Rieder, Saurav Chakraborty, Sandeep Goyal, and Donald Berndt on 'A Stratified Approach to Agent-Based Modelling: Unlocking Prediction in Non-Positivist Paradigms' criticises traditional agent-based modelling (ABM) for its lack of ontological stratification, which limits researchers' ability to capture complex causal mechanisms in IS research. Traditional ABM methods are criticised for focussing primarily on empirical stratification, separating micro- and macro-level interactions without considering ontological stratification. This oversight limits the ability to identify and codify generative mechanisms, resulting in models that fail to capture the full complexity of IS phenomena. The authors argue that without ontological stratification, ABMs cannot fully represent the causal powers and empirical events driving system dynamics. This gap hinders the development of robust theories that explain the emergence and evolution of complex social and technical systems. To address the above

limitations, the authors propose a stratified ABM approach that incorporates ontological stratification into the modeling process. This approach involves identifying generative mechanisms and codifying them into ABM simulations, providing a more comprehensive representation of causal powers and empirical events. The stratified ABM method includes distinct steps for identifying, codifying, and validating these mechanisms, bridging the gap between empirical observations and theoretical constructs. By integrating critical realism principles, the proposed method aims to enhance the explanatory power of ABMs and support the development of more accurate and robust IS theories. The stratified ABM approach is demonstrated through a simulation of account removal strategies on social networks. The authors detail the steps of identifying generative mechanisms, coding them into the ABM, and validating the simulation results. The case study illustrates how the stratified approach captures the dynamic and multi-level nature of social network interactions, providing insights into the effectiveness of different account removal strategies. The demonstration highlights the practical application of stratified ABM, showing its potential to improve predictive accuracy and theoretical robustness in IS research.

Finally, Amber Young, Syed Shuva, Tamara Roth, Yaping Zhu, and Alan Hevner contribute an article on ‘Ethical Design through Grounding and Evaluation (EDGE) of Information Systems to Achieve Social Impacts’. The authors point out the lack of integration of ethical considerations in traditional IS research methodologies, specifically in DSR. Traditional DSR methodologies often overlook ethical considerations during the design and evaluation processes, leading to interventions that may be effective technically but problematic ethically. By neglecting ethical implications, these methodologies risk creating systems that can have negative societal impacts. The lack of systematic incorporation of ethical theories in DSR limits the ability to develop IS interventions that are both technically effective and ethically sound. The authors introduce the ‘Ethical Design through Grounding and Evaluation’ (EDGE) approach, integrating ethical theories into the design and evaluation phases of IS research. EDGE involves identifying relevant ethical principles, incorporating them into the design of interventions, and continuously evaluating ethical performance. This approach ensures that ethical considerations are systematically addressed throughout the research process, leading to IS interventions that are both effective and socially responsible. The EDGE approach is demonstrated through a case study of a customer loyalty apps platform. The authors demonstrate how to integrate ethical principles into the design process such that ethical performance of IS artefacts can be continuously evaluated. The case study showcases the iterative nature of EDGE, highlighting how ethical considerations can be addressed at each research stage.

Towards a framework for next-generation IS research methods

Observation 1: Going beyond static research designs to capture dynamic and complex phenomena

In their aspirations, the papers accepted for this special issue are representative of all submissions that we received. At a higher level, the special issue papers address the following themes that can be seen – to some extent – as shortcomings of today’s methodological toolbox:

Temporal detail for improved understanding of dynamic phenomena. The authors of the accepted papers criticised existing research methods for not being able to deal with temporality and highly dynamic phenomena. Existing research methods are considered as being too static so that authors propose novel experimental paradigms such as MRTs (Pieper et al.) or novel ways of analysing and understanding temporal dynamics in (textual) trace data and CTC (Kishore et al.).

Capturing complex and granular patterns. The authors noted that existing research methods fail to capture complex and granular phenomena and corresponding data patterns. This is not only relating to highly granular data sources such as trace data (Kishore et al.), but also whether existing research methods are powerful enough to detect relevant patterns in data (Stoffels et al.) as well as generative patterns through which the phenomena under investigation are shaped (Rieder et al.). Against that backdrop, Stoffels et al. investigate the validity of XAI in terms of CTC and Rieder et al. focus on stratified ABM to investigate the explanatory power and accuracy of simulation models which address phenomena that preclude empirical research.

Computational theory construction. On a meta-level, the proposed methods attempt to study such temporal and complex phenomena by blurring ‘traditional’ boundaries between quantitative and qualitative research and building on ideas of CTC (Kishore et al.; Rieder et al.; Stoffels et al.). For instance, Kishore et al. focus on quantitatively analysing textual trace data and then engage in an interpretative theory construction process. Similarly, Rieder et al. introduce the notion of interpretative thinking into ABM.

Effective integration of ethical considerations. Finally, the paper of Young et al. takes a different stance and focuses on design-oriented research. The authors propose an extension of DSR by offering an approach to systematically integrating ethical considerations into the design and evaluation of IS artefacts.

Observation 2: GAI as (potential) productivity enhancer in research

Complementing the extension of today’s research methods to better understand dynamic and complex phenomena and

integrating ethical considerations more effectively offered by the special issue articles, we see wide-spread adoption of GAI. Consequently, academia has started an ample debate about how such tools could and should be used in academic research (Bail, 2024; Davison et al., 2024; Schlagwein and Willcocks, 2023; Stokel-Walker and Van Noorden, 2023; Van Dis et al., 2023). For instance, Susarla et al. (2023) discuss the potentials of GAI as a tool for conducting research. They highlight that GAI might aid researchers in finding ideas for novel research projects in terms of a sounding board, might facilitate the collection and analysis of data, as well as help the actual writing process. Generally, GAI might shorten the time-to-publication by making researchers more productive and could help democratise science to make it more equitable and approachable (Stokel-Walker and Van Noorden, 2023; Susarla et al., 2023).

Although research generally sees great potential for GAI, researchers also outline substantial challenges including the truthfulness of obtained outputs and fabricated answers, ethical questions concerning the transparency and integrity of the research process, or plagiarism and copyright infringements (Bail, 2024; Davison et al., 2024; Grimes et al., 2023; Sabherwal and Grover, 2024; Schlagwein and Willcocks, 2023; Stokel-Walker and Van Noorden, 2023; Susarla et al., 2023; Van Dis et al., 2023). For instance, Schlagwein and Willcocks (2023) discuss the ethical implications of using GAI in research highlighting that the lack of consensus of how GAI should be used might create ethical nonchalance and a deficit in social responsibility. In sum, the general argument of these articles is that GAI is not yet able to produce high-quality research on its own (yet), although it has considerable potential to support researchers along the processes of knowledge synthesis, creation, evaluation, and translation (Grimes et al., 2023; Susarla et al. 2023). Consequently, the prevailing metaphor for characterising this joint research process is the one of ‘human–AI collaboration’. For instance, Sabherwal and Grover (2024) highlight the notion of augmentation meaning that GAI can improve research productivity by delegating certain support tasks along the research process to GAI models instead of fully automating the entire job of doing research.

An organising framework of next-generation IS research methods

In Figure 1, we attempt to describe the changing nature of future research methods in the field of IS. The accepted papers of the special issue emphasise a better understanding of highly complex and dynamic phenomena by suggesting novel approaches to CTC. The ongoing proliferation of GAI is likely to accelerate this trajectory. Schlagwein and Willcocks (2023) mentioned that the *Journal of Information Technology* allows the usage of GAI tools if researchers

find productive and legitimate ways of using it. However, what is this productive and legitimate way of using GAI? We intend to contribute to illuminating this path and to extend the discussion on (1) what kind of GAI-enabled research support is legitimate in the paradigms of quantitative, qualitative, and engaged research, (2) how occurring challenges can be effectively mitigated, and (3) how the role of researchers might potentially change if research tasks are systematically delegated to GAI models. In the following sections, we attempt to answer these questions.

Generative AI as support for existing research methods

Potential for generative AI-enabled research

We highlight illustrative applications of GAI for qualitative, quantitative, and engaged research. We acknowledge that these applications might not be exhaustive and that certain GAI potentials that we describe in one of the three paradigms might also be applied in the other ones. Following Grimes et al. (2023), we decompose the scholarly process into (1) knowledge synthesis, (2) knowledge development and evaluation, and (3) knowledge translation. Across these three stages, GAI empowers researchers to achieve distinct levels of ‘know-what’ (awareness of data, information, concepts, and theories) and ‘know-how’ (procedural knowledge for conducting research). We think that the potentials of GAI in the phases of knowledge synthesis and translation are uniform across the three research paradigms, while there are specific differences in the phases of knowledge generation and evaluation. Table 1 highlights these communalities and differences which we describe below.

Knowledge synthesis

The process of conducting research begins with the ‘know-what’ stage, where researchers start with a broad area of interest and then narrow it down to a specific research question. This often involves a thorough review of existing literature and the development of hypotheses based on theory. This stage involves identifying what is already known about the topic and what gaps exist in the current knowledge. Thus, it may help to generate ideas for novel research projects.

Before the emergence of GAI, researchers had to manually review, summarise, and synthesise extensive bodies of literature to identify potential research gaps. GAI can expedite literature reviews (Ngwenyama and Rowe, 2024; Tinglehoff et al., 2024) and enhance our ability to leap across literatures and share/recombine diverse resources, facilitating and expediting novel and multi-disciplinary research (Kulkarni et al. 2024). For instance, researchers

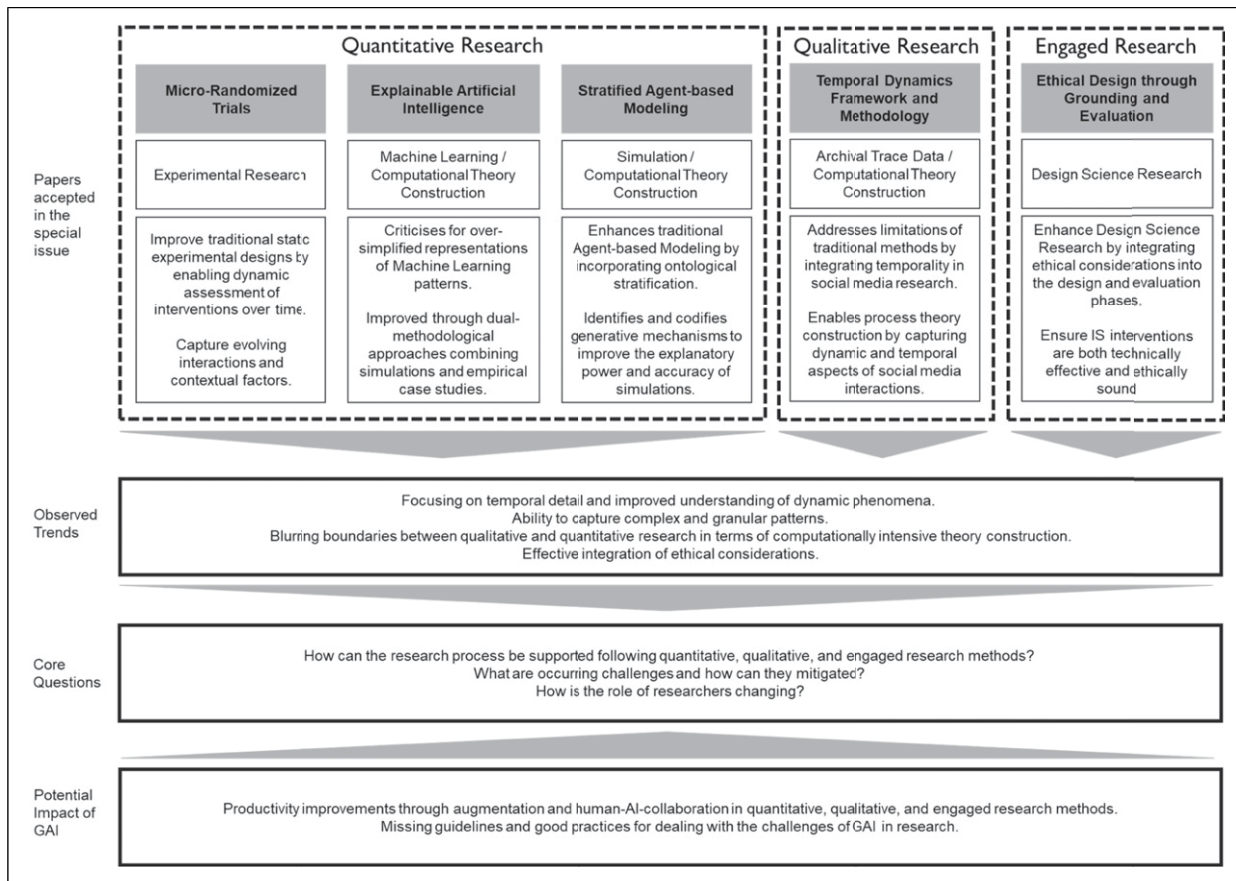


Figure 1. Framework for novel methodologies for future information systems research.

sometimes include appendix tables that summarise literature, listing the objectives of studies, theories adopted, samples, and major findings. While traditional tools like search engines and *Google Scholar* provide some assistance, the process is labour-intensive. GAI can synthesise large volumes of literature early on, conduct topic modelling to discern key themes (Gupta et al., 2024), and enable researchers to pinpoint gaps (Haman and Školník, 2023). GAI can be interrogated to synthesise literatures to identify underlying assumptions, limitations, and boundary conditions to inform quantitative research as well as harvest insights from kernel theories toward design science research (Jarvenpaa and Klein, 2024). Fine-tuned GAI models are being used to generate testable hypotheses from corpora of social psychology journal articles and preprints, with novelty scores that far surpass those generated by humans (Banker et al., 2024). The wealth of original hypotheses thus generated enables the researcher to critically assess them for worthiness for further empirical attention (Berger, 2024). Additionally, AI-powered citation analysis can reveal the context and impact of cited works, allowing researchers to gauge the quality and relevance of identified literature based on their citation influence (Fok et al., 2024).

Beyond the context of academic research, *Iris.ai* performs frequency analysis on text to identify and understand important words in context. It reads *TED* talk transcripts, clusters words into contextually similar groups using ML and organises concepts into flexible hierarchies. These are then presented through Voronoi treemaps for clear visualisation, showcasing insights from *TED* talks. This technology could also be applied to the literature review process to synthesise large volumes of journal publication texts, identify key themes, and highlight potential research gaps, aiding researchers in making informed decisions for their studies.

To understand the impact of each research publication, AI tools like *scite.ai* with the Smart Citations feature discover and evaluate scientific articles (Nicholson et al., 2021). This feature shows how publications are cited and classifies citations as supporting or contrasting evidence. *Scite.ai* and other similar AI-powered citation analysis highlight key contributions and ongoing debates, helping researchers assess publications with significant impact and credibility. The citation information also helps researchers determine whether a topic is saturated or still has high potential. However, recent research indicates that clustering

Table 1. Application potentials of GAI in qualitative, quantitative, and engaged research.

Research method	Knowledge synthesis	Knowledge development and evaluation	Knowledge translation
Qualitative research	<ul style="list-style-type: none"> Automated literature review and synthesis Topic modelling for key themes Identifying research gaps and ideas GAI-enabled citation analysis Discover and evaluate scientific articles 	<ul style="list-style-type: none"> Pre-processing qualitative data (note taking and transcription) Generating interview guidelines Conducting interviews at scale with conversational agents Enhanced qualitative data analysis (e.g. coding and theme identification) Combining GAI with traditional qualitative data analysis Evaluating validity and reliability of qualitative data analysis Evaluate coherence of research findings 	<ul style="list-style-type: none"> Challenge research findings from diverse theoretical and methodological backgrounds Summarising key insights for diverse audiences Visualise research findings
Quantitative research		<ul style="list-style-type: none"> Pre-processing and cleaning data sets Generating survey items and refining instruments Simulating datasets for pilot testing Generating synthetic data Contextualizing survey instruments for different settings Creating experimental materials Assisting in code creation for analysis Running and interpreting statistical analyses 	
Engaged research		<ul style="list-style-type: none"> Extracting user needs and requirements from unstructured data Simulate individuals from the problem domain for understanding problem with the status quo and latent needs Deriving potential solution ideas and problem solutions Evaluate design requirements and design choices Rapid prototyping of design options Coding support in implementing specific IS artefacts Creation of synthetic data to develop and test IS artefacts 	

Research activities in the knowledge development and evaluation phase might also be applied in other research methods beyond our categorization because corresponding research activities are partially overlapping. We put them into the 'most appropriate' category.

and visualisation based on GAI lack theory-guiding principles, and researcher intervention and improvement are deemed necessary (Haque, 2024).

Knowledge development and evaluation

Quantitative research. Knowledge development involves data collection, modelling, analysis, and crafting clear arguments. GAI lowers methodological barriers and speeds up tasks like research and instrument design (Grimes et al., 2023). An important 'know-how' aspect of quantitative research is the development and refinement of survey instruments. Given the fluid nature and broadened scope of IS

research, researchers often need to develop new instruments to capture new concepts. Traditionally, IS researchers were expected to adapt or use verified and established instruments extracted from the literature, or they conducted rigorous procedures to develop and validate new instruments. However, this manual development of new instruments is labour-intensive and time-demanding. When researchers propose survey items in an ad hoc manner without validation, journal reviewers sometimes question the quality of these measures.

GAI, particularly LLMs, can generate new survey items due to their extensive training on massive amounts of text data from diverse sources such as books, news articles, web

pages, and social media posts. This broad training enables LLMs to understand various contexts, topics, and nuances in language. Advances in model building and training have led to the emergence of GAI models that can be adapted to a wide range of tasks, including those they were not originally trained for, such as generating survey items. The ability to comprehend and generate human-like text allows LLMs to create relevant, clear, and varied survey questions that align with specific research objectives, making them a valuable tool in survey development. Researchers may use GAI to fine-tune questionnaire items, ensuring comprehensive coverage (Salah et al., 2024).

Additionally, researchers can generate a simulated dataset based on the distribution of potential participants' demographics and their responses to the questionnaire items. In medical research, synthetic datasets are commonly used to train deep learning models (e.g. Umer and Adnan, 2024). In IS, researchers can simulate demographic data using appropriate statistical distributions; for instance, age can be modelled by a normal distribution with a realistic range, and income can be modelled by a log-normal distribution. Subsequently, researchers can simulate survey responses based on the constructed sample's demographic profiles using randomization to ensure realistic distributions. Using these simulated datasets as a pilot run, researchers can apply techniques such as factor and reliability analysis to evaluate and refine the survey items. Likewise, LLMs can advise researchers on additional variables to include in their studies (Susarla et al., 2023). By analysing vast amounts of text data from diverse sources, LLMs can provide insights into commonly studied variables and emerging trends within specific fields.

With a set of developed items, researchers leverage GAI to contextualise survey instruments to, for instance, industry culture, regional contexts, and user groups by translating and adapting survey items to ensure their relevance and meaningfulness across different settings. This also enables researchers to efficiently develop different versions of survey instruments for conducting comparative analyses, leading to higher practical contributions. However, IS researchers are encouraged to develop effective validation approaches to assess a survey instrument's validity, comprehensiveness, and completeness. Validation methods may include pilot testing, factor analysis, reliability testing, and content validity assessments involving experts from relevant fields. These steps are crucial to ensure that survey instruments accurately capture the intended constructs and are applicable across various contexts.

Beyond the creation of scales for survey-based research, we also see a series of opportunities for experimental research. For instance, GAI facilitates the creation of highly realistic experimental materials such as text, images, and videos (Rossi et al., 2024). Models like *GPT* can generate human-like text, enabling the production of realistic social

media posts, emails, and other textual content for experiments. Similarly, tools like *DALL-E* and *Stable Diffusion* can produce photorealistic images, making stimuli nearly indistinguishable from real-world items. This enhancement in realism contributes to more ecologically valid experiments, increasing the external validity of research findings (Rossi et al., 2024).

During the evaluation phase, GAI can be employed to analyse collected data with increased efficiency and rigour (Grimes et al., 2023). During this process, GAI might be helpful in pre-processing and cleaning data. For instance, Susarla et al. (2023) highlight the potentials of GAI models for pattern matching. For instance, when combining data sets from diverse sources, researchers have to match records which usually involve manual labour (e.g. variations of company names). GAI might reflect an effective solution to speed up such activities (Susarla et al., 2023). Beyond that, AI tools can identify patterns, inconsistencies, and insights within large datasets, offering robust statistical analyses. Well-known platforms including *Azure Machine Learning*, *Google Cloud AI Platform*, and *Amazon Web Services* provide a comprehensive suite of services to build, deploy, and manage ML models.

Currently, there is no full automation of data analyses; researchers still hold unique value in this process. Their unique experiences and accumulated knowledge, referred to as 'bounded rationality' (Sabherwal and Grover 2024: 19), enable them to navigate complex datasets and derive meaningful interpretations. This expertise provides researchers with a competitive advantage, as they can contextualise findings, make informed decisions and apply ethical considerations that GAI alone may not fully address. However, GAI tools such as *GitHub Copilot* help in creating the code for statistical analysis tools such as *R* or *Python*, running the actual analysis, or helping to interpret the obtained results (Susarla et al., 2023).

As GAI data analyses become more mature, the process may become more fully automated. However, this advancement brings concerns about the potential for errors within these complex models. Identifying mistakes and ensuring the accuracy of analyses performed by GAI will become increasingly challenging. Human oversight will be crucial to validate results and maintain the integrity of data interpretations. The trajectory towards 'collective rationality' (Sabherwal and Grover 2024: 19), while beneficial in creating a vast knowledge base, could also dilute individual expertise and threaten human uniqueness and dignity. Ensuring equitable distribution of the benefits from leveraging this collective knowledge will be essential to avoid centralization of power and maintain the value of human contributions in the data analysis process. It is imperative that researchers continue to check the outputs of GAI to prevent the spread of misinformation and the creation of distorted digital realities (Sabherwal and Grover, 2024).

Qualitative research. Qualitative methods comprise a variety of genres. Across these genres, the nature of data and the desired knowledge claims differ as do the evaluation criteria. Understandably therefore, how GAI should be used and its use judged should be contextualised to the specific genre within which it is being applied. The qualitative methods landscape is further complicated by researchers blending different genres within a single study. For example, [Durcikova et al. \(2024\)](#) began with open coding of interview transcripts, that is, an inductive approach, followed it up with content analysis, that is, a deductive approach, of the transcripts and then subjected the coded categories to further abductive processes to develop a theory of how cybersecurity units become and remain empowered. While earlier generations of AI have fostered innovations in qualitative methods, GAI promises qualitative research ‘at scale’, permitting the researcher to perform tasks associated with conventional qualitative methods more expeditiously and availing of larger data sets. Unlike other paradigms, the knowledge processes entailed in qualitative research are iterative, rather than linear ([Locke et al., 2022](#)). Rather than progressions from knowledge synthesis to development to evaluation to translation, development and evaluation processes must feed each other.

Across the different qualitative genres, GAI is useful in data collection and pre-processing. It is being used to develop and refine interview protocols ([Parker et al., 2023](#)). Conversational agents are being used to conduct interviews in HR and customer service processes ([Nawaz and Gomes, 2019](#); [Sidaoui et al., 2020](#)). They also are being used for conducting scholarly interviews ([Chopra and Haaland, 2023](#); [Villalba et al., 2023](#)), promoting the notion of qualitative research ‘at scale’. LLMs are useful for collecting the unstructured data – such as email, social media posts, and e-forum discussions – that form the basis of qualitative research ([Kulkarni et al., 2024](#)). In assisting with data collection though, LLM-based tools may lack the ability to contextualise the data ([Kulkarni et al., 2024](#)), and contextualization is essential to qualitative inquiry ([Cecez-Kecmanovic et al., 2020](#)).

GAI tools also help with the pre-processing of qualitative data. LLMs assist with pre-processing textual data ([Kulkarni et al., 2024](#)). Apps such as *notta* and *Samsung Galaxy AI* offer live translation of interviews. *Otter.AI*, *notta*, and others facilitate note taking and automate transcription. Nonetheless, in our research, we have found it essential for a member of the research team or a research assistant to check the automated transcriptions against interview recordings as transcriptions often distort unusual words or accents.

Different algorithmic techniques are available to assist with deductive versus inductive/abductive analyses of qualitative data. Researchers using deductive methods on qualitative data typically engage in content analysis. For

such researchers, algorithmic techniques can expedite content analysis. For example, [Mousavi and Gu \(2024\)](#) applied three algorithmic techniques – a dictionary-based approach, an LSTM approach, and a *BERT* model – to code tweets for the presence of resilience messaging. Recently, [Pattyn \(2024\)](#) experimented with deploying GAI in content analysis. Specifically, he compared coding by novice and expert researchers and by *Open AI’s ChatGPT* and *Google’s Bard* (now called *Gemini*) in two independent sessions. He found that using *ChatGPT* did result in efficiency and inter-rater reliability (Cohen’s kappa) gains, but the highest gains in inter-rater reliability were for hybrid teams composed of expert coders and *ChatGPT*.

Researchers have made extensive use of traditional algorithms such as topic modelling for document summarization and open-coding of large corpora in inductive/abductive work ([Croideu and Kim, 2018](#); [Miranda et al., 2022b](#)). They have also applied algorithms such as multi-dimensional scaling to these first-order codes for further thematic analysis ([Miranda et al., 2022b](#)). GAI makes computational tasks such as topic modelling more accessible to novice programmers. For example, [Mansurova \(2023\)](#) describes using *ChatGPT* for *BERT* topic modelling. Researchers are also reporting being able to derive greater insight from their qualitative data when combining GAI with use of traditional qualitative data analysis tools such as *MAXQDA* ([Perkins and Roe, 2024](#)). For example, [Sinha et al. \(2024\)](#) experimented with having *Microsoft Copilot* perform line-by-line coding, along with two human coders using *MAXQDA* in a constructivist grounded theory project. They iteratively combined the three sets of codes into the knowledge artefacts that culled and represented their theoretical insights from the data. They reported that *Microsoft Copilot* was useful in identifying data segments that the human coders missed. However, they noted that it applied a coarser grain in coding, that is, abstracting codes from bigger chunks of text than did the human coders, relative to the more nuanced coding performed by the human coders. Finally, GAI can now directly perform the kinds of document summarization, text classification, and sentiment analysis tasks that were previously computationally mediated, further broadening access to these analytic processes ([Zhang et al., 2023](#)).

Intermediate and final theoretical insights from inductive qualitative analysis are often visually represented; in fact, a visual representation is expected for grounded theory insights ([Creswell, 2007](#)). Our experience with using such tools for visualisation is that they do not produce usable visuals but may inspire novel visualisations that the researcher may otherwise not have considered.

In evaluating the knowledge produced by positivist qualitative research, the criteria for quantitative behavioural research pertain, that is, reliability and validity. Coherence of research findings with an existing body of knowledge is

also essential, especially for inductive and abductive knowledge production (Miranda et al., 2022a). As Kulkarni et al. (2024, 209) noted: ‘The theoretical contributions we make are not so much based on filling objectively discoverable ‘gaps in the body of knowledge’ but rather narrative constructions in an evolving, intertextual discourse’. For this, speciality literature review tools such as *elicit.com*, *scite.ai*, and *consensus.app* may prove useful (Tingelhoff et al., 2024). Above all, however, surprise and novelty are the holy grail for inductive/abductive qualitative research (Sætre and Van de Ven, 2021). GAI may help researchers quickly home in on what is surprising by identifying infrequent words or breaks in empirical patterns (Kulkarni et al. 2024).

Engaged research. Engaged research methods such as design science and action research focus on solving prevailing individual, organisational, or societal problems. As opposed to empirical qualitative and quantitative research, their primary goal is not to construct or test theory. By contrast, they are focussing on creating useful and innovative problem solutions. While DSR focuses on the purposeful design and creation of (technological) artefacts (Peffers et al., 2007; Tuunanen et al., 2024), action researchers rather focus on inducing and observing change in social systems to address the focal problem (Baskerville and Wood-Harper, 1996). In this context, the induction of change can be considered as some sort of intervention or action-taking that may or may not involve the creation of a specific (technological) artefact.

Engaged research usually begins with developing a detailed understanding of the problem that researchers want to solve. In DSR, the first research phase usually refers to defining the specific research problem as well as developing a corresponding problem statement (Peffers et al., 2007; Tuunanen et al., 2024). Similarly, action research speaks of ‘diagnosing’ to analyse the current status quo. Either way, in both approaches this process includes atomizing the problem conceptually so that the envisioned solution can capture the problem’s complexity (Peffers et al., 2007). Against that backdrop, GAI may be effective for exploring the initial problem space (Bouschery et al., 2023). For instance, Giunti and Doherty (2024) describe an automated approach to crawl app reviews with GAI, while Bouschery et al. (2023) focus on discussing how GAI may help extract insights from such a large corpus of unstructured textual data by summarising this textual data to distil user problems effectively so that researchers can speed up knowledge extraction (Bouschery et al., 2023). Similarly, GAI could be used to simulate potential users or individuals from the problem domain which can be interviewed to better understand problems with existing solutions and the status quo to identify latent needs (Ataei et al., 2025). During this process, GAI models might be asked to simulate interview

partners, for example, taking the role of an ‘average user’ or more advanced ‘experts’ (Ataei et al., 2025).

The second step in DSR usually refers to ‘objectives and requirements definition’ in which the problem statement is translated into objectives and design requirements (Peffers et al., 2007; Tuunanen et al., 2024). In a similar vein, action research refers to ‘action planning’. During these phases process, GAI could be used to explore the solution space (Bouschery et al., 2023), for example, by generating potential solution ideas to a given problem statement, giving feedback to solution ideas that have been generated by human researchers (Gordetzki et al., 2023), or extending simulated interviews to elicit requirements (Ataei et al., 2025). Following this line of reasoning, Memmert et al. (2023) show that GAI-enabled conversational agents provide helpful inspirational suggestions for design requirements by asking open-ended questions and encouraging the elaboration of initial ideas in DSR projects. More specifically, GAI could help challenge the elicited design requirements in terms of applicability to the problem statement or coherence to avoid contradictions (Tuunanen et al., 2024).

Extending this line of thought, GAI might also help researchers in the ‘design and development’ of artefacts and/or interventions to address the focal problem. For example, in DSR, GAI may support researchers in creating design principles (Memmert et al., 2023) or other parts of the solution design. Following the notion of rapid prototyping (Bilgram and Laarmann, 2023), GAI may support researchers in creating alternative design ideas so that a broader solution scope can be explored. For instance, GAI might help to create alternative design mock-ups of a user interface or suggest different options to implement certain design features of a solution design. During this process, GAI might also help evaluate these design choices, for example, in terms of feasibility or consistency (Tuunanen et al., 2024), and selecting the most appropriate ones for implementation.

Finally, GAI might also support researchers in the implementation of the design, for example, by providing low code frameworks or general coding help for implementing the actual IS artefact. This is not only addressing the coding parts, but also the creation of synthetic development and test data. For instance, in many contexts it might be difficult to obtain sufficient real-world data to develop potential artefacts and/or problem solutions. In such contexts, GAI might also be used to generate synthetic data that closely resembles real data, which is valuable when real data is scarce, expensive, or sensitive (Rossi et al., 2024). For example, synthetic patient data in healthcare can train diagnostic models without compromising patient privacy. This approach allows researchers to test hypotheses and conduct simulations without the ethical and logistical challenges associated with real data. Similarly, GAI could be used to simulate users with many different traits so that the functionality of the system can be

tested in a more fine-grained and effective fashion. While synthetic data offers numerous benefits, its reliability and validity must be rigorously evaluated. Synthetic data may not perfectly capture the nuances and variability of real data, potentially leading to misleading conclusions (Rossi et al., 2024). Researchers must develop robust validation frameworks to ensure that synthetic data accurately represents the phenomena under study. This involves comparing synthetic data with real-world data to assess its accuracy and generalizability (Rossi et al., 2024).

Design and action researchers have adopted a broad variety of methodologies for demonstrating and evaluating the usefulness of their artefacts and interventions including qualitative (e.g. expert or user interviews) and quantitative approaches (e.g. surveys or experiments) (Tuunanen et al., 2024). We refer to the respective sections of this editorial to showcase application potentials for knowledge evaluation in engaged research methods.

Knowledge translation

In the final knowledge translation phase, researchers aim to communicate the most valuable aspects of their projects to diverse audience groups. GAI can significantly aid in converting research findings into accessible and impactful knowledge tailored for different languages and audiences. For instance, *ChatGPT* can assist in drafting research reports, creating visualisations for complex ideas, and translating findings into multiple languages, thus broadening the reach and impact of the research. Currently, *ChatGPT* can translate written content into fifteen languages, making research results more accessible to a global audience. Developed by Google, *notebookLM* can generate podcast like audio overviews for research papers, enhancing the accessibility and comprehensibility of research findings. *Zoho ShowTime* leverages AI to assist in creating presentation materials. Further, the potentials of GAI in terms of concept and text simplification might help researchers to more effectively convey the outcomes of their research to the general public (Cardon and Bibal, 2023). Potentially, GAI might also help to challenge a manuscript from different theoretical and methodological perspectives, helping researchers to identify and address potential misunderstandings and improve their work (Susarla et al., 2023). Similarly, tools like *TableauGPT* can generate interactive charts and graphs from textual or numerical data, making complex data more understandable. These capabilities ensure that complex research findings are effectively communicated to both academic and non-academic audiences, bridging the gap between theory and practice, ensuring that valuable insights reach and benefit a broader audience through various channels.

Beyond these important activities, the process of knowledge translation does also relate to refine empirical results of knowledge evaluation endeavours into theoretical

contributions (Grimes et al., 2023). Although we see that GAI might potentially be capable in crafting theoretical contributions, we think that GAI is subject to certain challenges that might hamper GAI's potential to this core discipline of knowledge translation.

Challenges of generative AI application in research

The integrity of scholarly research relies on methodological transparency and reliability, and large-scale implementation of LLMs raises concerns about these very aspects. The generation of false yet plausible and contextually relevant information by AI can mislead researchers. Furthermore, 'deep research fakes', where GAI is deliberately used to fabricate and manipulate data to deceive the scholarly community, represent a serious threat. In Table 2, we outline key challenges of applying GAI in research.

Authenticity of research and researcher's voice

As technology progresses and plays a larger role in human-centred scholarship, concerns about the authenticity of scholarly work increase. The involvement of GAI in research necessitates careful consideration of the circumstances under which human judgement is crucial to ensure authenticity (Fui-Hoon Nah et al., 2023). Major publishers like *Springer-Nature* and *Elsevier* have updated their authorship policies, clarifying that AI tools, including *ChatGPT*, cannot be credited as authors. These tools can enhance language and readability but cannot replace the intellectual contributions of human researchers. Consequently, *ChatGPT* should be seen as a complementary tool to help polish the writing of research ideas rather than a substitute for traditional academic practices. Its role is best suited for assisting researchers with preliminary tasks and enhancing efficiency, but not for tasks requiring critical thinking, originality, and accountability. Researchers must use GAI ethically, acknowledging its contributions appropriately and ensuring the integrity and validity of their work.

GAI also draws concerns about whose voice is represented. For example, Saku Mantere and Eero Vaara write: 'ChatGPT's voice is essentially a corporate voice that is governed by, rather than democratically chosen values, principles of mitigating corporate risk' (Kulkarni et al., 2024: 210) remaining purely descriptive and displaying a 'moral indifference born of unintelligence'.

Replicability and transparency of the research process

The current generation of GAI models is of probabilistic nature so that the same input (prompt) may lead to varying

Table 2. Key challenges for applying generative AI in research.

Theme	Key points
Authenticity of research and researcher's voice	<ul style="list-style-type: none"> • Human judgement remains crucial for authenticity • Publishers prohibit GAI tools, for example, ChatGPT, from being credited as authors • GAI enhances language but cannot replace critical thinking, originality, and accountability • Ethical usage and proper acknowledgement are essential
Replicability and transparency of research process	<ul style="list-style-type: none"> • GAI's probabilistic nature challenges replicability of the research process • Constant development of GAI requires frequent adaptation • Performance varies between GAI models • Proprietary nature of GAI hinders transparency • Privacy issues may arise when using GAI in data analysis
Algorithmic isomorphism and originality	<ul style="list-style-type: none"> • GAI struggles to produce novel ideas and genuine research gaps • Outputs may reflect reiteration of accepted wisdom • Public LLMs lack specificity due to reliance on non-academic data
Evaluation of output quality and content hacking	<ul style="list-style-type: none"> • GAI models hallucinate and may generate outputs with no basis in fact • Challenges in truthfulness and evidence verification persist • Risks of 'content hacking' akin to 'p-hacking' • Validation of GAI output is necessary but lacks standardized metrics • Multi-agent systems such as 'LLM-as-a-judge' show promise but raise questions regarding validity and reliability
Perpetuation of bias	<ul style="list-style-type: none"> • Tacit depiction of social sciences as non-scientific • Variation in usefulness, level of detail, and certainty of knowledge claims relative to user profile information and query history
Undermining the academic enterprise	<ul style="list-style-type: none"> • Overloading the volunteer peer-review process • Lowering our thresholds for sharing private information and copyrighted products and vulnerability to prompt injection hacks • Eroding critical thinking skills necessary for good science

results and unpredictable behaviour of a GAI model (Bail, 2024). While this behaviour can be considered as part of the experience of maintaining a 'real' conversation or interaction with a GAI model and might even be beneficial in gaining inspiration in the early phases of a research project, the lack of replicability is at odds with one of the core pillars of science (Bail, 2024). Bail (2024) highlights two replicability challenges. First, (commercial) GAI models are under constant technological development so that their results might vary over time (Bail, 2024). Consequently, prompts might require constant fine-tuning and adaptation to deliver similar results which complicates the reproducibility of research. Second, the performance of different GAI models may vary drastically across distinct tasks making it difficult for researchers to choose an 'appropriate' GAI model. For instance, Kim et al. (2024) compare the ability of *ChatGPT* and *Bard* in writing paper abstracts. While *ChatGPT* did a better job in corresponding to a journal's formatting guidelines, *Bard* performed better in meeting the abstract's expected word counts. However, for both models, humans were not able to accurately distinguish between human and GAI written abstracts (Kim et al., 2024).

Thus, ensuring replicability of GAI-created research results might be challenging despite initial options to control the behaviour of GAI models. For instance, some

GAI models offer parameters to control the variability of the output, for example, the temperature parameter in *ChatGPT*. Alternatively, initial programming frameworks such *DSPy* have been developed which abstract from manually created prompts to make the response of GAI models more reliable (Khattab et al., 2023). Against this backdrop, researchers are encouraged to meticulously document their research efforts (Susarla et al., 2023). Disclosure of tools such as software libraries to ensure transparency of the research process and informed consent of respondents during data collection (Davison et al. 2024) and beyond (Schlagwein and Willcocks, 2023) is also a key recommendation.

All these challenges also refer to the transparency of the research process as such. First, today's most powerful and accessible GAI models are usually of proprietary nature, for example, *OpenAI's* family of *GPT* models. As such it is not yet clear which training data has exactly been used and how the models have been trained (Susarla et al. 2023). In some cases, even the architecture of these models may not be known. Extending this line of reasoning, the interpretability of the output produced by GAI models might hamper the transparency of the research process. GAI models learn a probability distribution over a high-dimensional space so that it might be difficult to understand for researchers why a specific suggestion is made, while others are not mentioned (Susarla et al., 2023). As a consequence, the reasoning

behind such challenges might remain opaque (Schlagwein and Willcocks, 2023). Second, Davison et al. (2024) argue that the usage of GAI tools for data analysis might hamper the transparency of the research process in terms of data analysis. They argue that the respondents' privacy rights because their data might also be shared with a third party, that is, a (commercial) GAI provider in terms of GAI-enabled research tools or cloud-based GAI models. This might have important consequences in terms of obtaining informed consent of respondents, ensuring adherence to non-disclosure-agreements, or obtaining approval of institutional review boards.

Extending this line of thought, Van Dis et al. (2023) encourage the development and use of open-source GAI models that can potentially be run on local infrastructure. While such arrangements are preferably in terms of privacy and transparency, the available options of commercial vendors could be more advanced in terms of truthfulness of the created output. Thus, researchers facing the question of selecting the 'most appropriate' GAI model might quickly enter a dilemma in which they might balance truthfulness of results against the transparency of the research process (Schlagwein and Willcocks, 2023).

Algorithmic isomorphism and originality of research

Because GAI models are synthetic and predictive, 'there is a genuine risk that the knowledge produced will be more a reiteration of accepted wisdom rather than genuinely new ideas or novel insights as findings become less and less disruptive over time' (Kulkarni et al., 2024: 210). It is notable that while GAI synthesises literature, it struggles with creating original research problem statements, identifying genuine research gaps (Rahman et al., 2023). This is especially a threat for knowledge discovery endeavours, in which novelty is key (Kulkarni et al. 2024). Evidence from research on the consequences of GAI use in other creative endeavours is likewise cautionary. For example, Chen and Chan (2024), investigating LLM use in writing advertising copy found using it as a sounding board enhanced output quality for non-experts; using it as a ghostwriter was in fact detrimental to output quality for experts. Likewise, Doshi and Hauser (2024) found that while GAI use by writers resulted in individual users authoring better stories, collectively, the narratives produced were less diverse. Further, as materials produced by GAI or in conjunction feed into LLMs as training data, the algorithmic recursion can lead to model collapse as the 'tails of the original content distribution disappear' (Shumailov et al., 2023: 1). Similarly, public LLMs are trained on publicly available knowledge which is usually of non-academic nature (e.g. Stokel-Walker and Van Noorden, 2023); as such, pedigreed knowledge, which is often copyrighted, is underrepresented in the repertoire of LLMs so that it may lack the level of

specificity that is required for academic research (Stokel-Walker and Van Noorden, 2023). This limits the usefulness of GAI for generating ideas that are truly leading edge as pedigreed knowledge will need to have crossed the scientist-lay person boundary before being available to public LLMs. Notably though, Banker et al. (2024) showed that by adjusting the 'temperature' parameter, GPT-3 could be tuned to heighten novelty in hypotheses advanced, albeit with an increased risk of generating hallucinations or nonsense.

Evaluation of output quality and content hacking

Today, it is well-known that GAI models tend to hallucinate and face challenges of answering factual questions (Fui-Hoon Nah et al., 2023; Schlagwein and Willcocks, 2023; Stokel-Walker and Van Noorden, 2023; Susarla et al., 2023). This might lead researchers to make wrong conclusions, produce erroneous research, and/or plagiarise existing research. As GAI models have been trained to create well-articulated, profound, and confident answers (Schlagwein and Willcocks, 2023), even if these answers might be wrong, researchers face an increased need for the ability to critically reflect on the quality of GAI's output. This challenge is not only relating to the truthfulness of the output, but also to the lack of evidence and potentially biased outputs (Fui-Hoon Nah et al., 2023; Memmert et al., 2023; Schlagwein and Willcocks, 2023). These biases can appear in the generated content, leading to skewed results or reinforcing harmful stereotypes. For example, image generation models might produce biased representations based on the demographic distribution of their training data. Similarly, it has already been shown that GAI models work better for western cultures than for eastern ones (Stokel-Walker and Van Noorden, 2023).

Addressing these biases requires careful consideration and implementation of fairness and bias mitigation strategies. While obtaining biased results is a general threat that is related to any kind of ML model, the scale and scope of how GAI models are used is aggravating this problem. Additionally, GAI sometimes generates fictitious references and researchers must exercise caution when using AI-generated content, ensuring thorough verification and cross-checking (Stokel-Walker and Van Noorden, 2023). Schlagwein and Willcocks (2023) highlight that researchers must take full responsibility for the generated output.

An additional GAI flaw noted by Chomsky et al. (2023) is that the algorithms – oriented toward description and prediction – can articulate statements about what *is* based on input data, but not what *is not*. The latter capacity for counterfactual reasoning is still uniquely human. Further, researchers are stating that potentially negative consequences of using GAI models in the research process cannot be foreseen (Schlagwein and Willcocks, 2023; Susarla et al.,

2023). The rise of ML/AI techniques have already challenged assumptions of quantitative research in which practices such as HARKing (Kerr 1998) – hypothesising after the results are known (aka p-hacking) – believed to lead to logically and scientifically flawed hypotheses, against concerns about stifling the advancement of knowledge (Pratt et al., 2019).

We share similar concerns in relation to using GAI in research. Practices such as prompt engineering and optimization have become common standards in the usage of GAI models. While prompt engineering reflects the systematic variation of instructions that are given to a GAI model to ensure the desired quality of the generated output (White et al., 2023), prompt optimization refers to the programmatic optimization of a given prompt (Khattab et al., 2023). Given a specific evaluation function, a computer program automatically updates the prompt until a certain quality threshold is reached. However, while these approaches might help make the results of GAI more reliable and predictable (Khattab et al., 2023; White et al., 2023), they might also increase novel challenges to legitimate research. Extending the line of argumentation that is prevalent to ‘p-hacking’, researchers may easily engage in a process of ‘content hacking’ in which they alter their prompts as long as they receive the results that they wish to obtain. Thus, we do not only encourage researchers to use GAI tools honestly (Stokel-Walker and Van Noorden, 2023), but also to carefully validate the outcome of GAI-enabled research.

The problem occurs because we currently do not have metrics for ensuring the validity and reliability for the output of GAI models. Thus, we believe that researchers must go beyond checking mere face validity when assessing the quality of a GAI model and engage in a rigorous validation process. Interestingly, first researchers have already started to automate such validation procedures following the ideas of multi-agent systems, that is, one GAI model is generating output while a second one is checking its quality (Zheng et al., 2024). However, as of today it remains an open question whether this ‘LLM-As-A-Judge’ technique produces valid and reliable assessments in an academic context – beyond all the ethical questions that this approach might invoke.

Perpetuation of bias

Social science researchers, particularly those from non-positivist traditions, often experience dismissive reactions and a bias against their work (e.g., Sarker et al., 2013). Research suggests that GAI perpetuates this bias. Specifically, Volk et al. (2024) found that *ChatGPT* manifested an overall tendency to view research as limited to disciplines from science, technology, engineering, and mathematics (STEM), largely excluding the social sciences. The

researchers characterized *ChatGPT*’s portrayal of science as a ‘positivist and empiricist’ endeavour (p. 8). Further, GAI responses are influenced by the user profile, the user’s prompt history – GAIs’ ‘long-term conversation memory’ (Goodin 2024), and how a prompt is phrased.

In setting up a *ChatGPT* profile, for example, a user may provide answers to questions about oneself such as: ‘Where are you based?’, ‘What do you do for work?’, ‘What are your interests and hobbies?’, and ‘What are some goals you have?’ They may also specify how they wish *ChatGPT* to respond – casually or formally, in detail or briefly or how they wish to be addressed. Jain and Jain (2024) found that responses to a prompt about pressing issues in neuroscience research were identified as a lack of technology to measure neural activity in the brain when the profile identified the interrogator as a scientist from Harvard University and a lack of funding and technology access when the profile identified the interrogator as a scientist from the Indian Institute of Technology. Even if users do not disclose such information in their profile, the GAI may be able to infer it from their prompt history (Jain and Jain 2024). Volk et al. (2024) found that while the substance of *ChatGPT* responses to scientific inquiries did not vary relative to user profiles, the responses did vary in their level of detail and supporting information provided, as well as in the certainty expressed. Specifically, responses to those who characterized themselves as ‘sciencephiles’ who trusted science, received responses reflecting a higher degree of certainty in the scientific knowledge reported than did others.

Jain and Jain (2024) further raised the concern that biases in the information used to train LLMs can be perpetuated through the scientific products that use outputs of those LLMs. Finally, biases can be introduced when scientists accept as fact GAI hallucinations, which are prone to occur with greater frequency when users juxtapose unrelated subjects in their query (Jain and Jain, 2024).

Undermining the academic enterprise

From the discussion above, it becomes clear that GAI can expedite many scientific processes. As a generative, rather than evaluative, technology though, GAI cannot help academics vet the scientific products generated though. The first threat to the academic enterprise, which has hitherto relied on a volunteer peer-review model for assessing scientific quality, will be the overloading of the peer-review model, which is already severely taxed. As the process efficiencies wrought by GAI translate into an explosion in the volume of journal submissions, the volunteer peer-review model will be unable to sustain the knowledge assessment process.

A second possible threat to the academic enterprise is in regard to privacy and ownership of intellectual property. As our engagement with GAI tools grows, our mindfulness and

intentionality in revealing information about ourselves and sharing others' intellectual property will likely diminish. As personal information and intellectual property shared feeds back into training underlying LLM models, the information and intellectual property move into the public sphere. Sabherwal and Grover (2024) bemoan the reliance of GAI on the subset of human experiences that have been publicly available to train the underlying LLMs. However, broadening the LLMs' training data via our GAI interactions that feed back into the models can also have insidious consequences. It can cause privacy violations, which may permit novel science-of-science studies but also manifest in undesirable profiling of individual researchers or categories of researchers, and violations of copyright. Further, a cybersecurity researcher has highlighted the ability of hackers to inject untrusted content from their personal emails, blogs, and other documents, which are then used to train the LLM, and then hack targeted GAI users the GAI's long-term conversation memory (Goodin, 2024). The researcher was able to trick *ChatGPT* into sharing a targeted *ChatGPT* user's entire prompt history and *ChatGPT* responses, further challenging individual privacy and intellectual property rights.

Finally, our progressive reliance on GAI risks limiting development of the critical thinking skills so essential to good scientific research. An uncritical reliance on GAI for literature reviews, hypothesis generation, data analysis, and writing erodes our ability to undertake those tasks. If this happens at scale, an absence of novel human-produced research on which to continue training LLMs will result in the predicted collapse of the LLMs (Shumailov et al., 2023), with researchers insufficiently trained to perform the scientific reasoning tasks for which they had come to depend on GAI.

Implications for the role of researchers

Sabherwal and Grover (2024) argued for a multiplier effect such that GAI increases the value of human capital. Researchers not using GAI will be at a competitive disadvantage relative to researchers using GAI (Sabherwal and Grover, 2024). It is therefore imperative that we develop our understanding of GAI capabilities and learn to deploy them effectively.

To understand the implications of GAI use in our research, we need to consider the nature of algorithmic agency and the relationship between algorithmic agents and humans. Algorithms have acquired a level of autonomy permitting us to view them as agents or actors, rather than simply as passive tools (Baird and Maruping, 2021; Kulkarni et al., 2024). As agential artefacts, researchers have begun to theorise the nature of algorithmic agency and the nature of human relationships with algorithms. The relationship between human and algorithmic agents has

conventionally been described in terms of automation (no human agency) or augmentation (algorithmically enhanced human agency) (Raisch and Krakowski, 2021). In contrast to this binary view of algorithmic and human agency, Baird and Maruping (2021) introduced the idea that humans delegate tasks to algorithms and the delegation is based on the respective endowments, preferences, and roles of the human and algorithmic agents. In Table 3, we consider the roles GAI may play in the research process based on the relative human and GAI endowments and human preferences.

When the GAI is cast in the role of *assistant*, the human delegator has strong a priori knowledge about the task. The objective of the delegation is process efficiency. The human has the capability to determine whether the output is acceptable and therefore to exercise output control. For example, a researcher assigning a deductive coding task to the GAI can compute inter-rater reliability of the GAI's codes with those of a human coder. If inter-rater reliabilities are low, the researcher can then feedback human codes to train the GAI to perform better. Similarly, researchers might apply GAI models to support them in creating survey scales or deriving potential requirements for a solution from a large corpus of user-generated content in a DSR project. In all these examples in which GAI models are used as an assistant, the critical assessment of the generated output is key. This also has implications for the role of researchers, potentially shifting it from the production to the evaluation of original content.

When the GAI is cast in the role of *teacher*, the objective of the human delegator is learning. The endowments of the GAI exceed that of the human delegator, who therefore lacks the knowledge to assess the quality of the GAI output. For example, a novice researcher wishing to perform topic modelling may request the code from a GAI. The accessibility of GAI tools lowers the barriers to entry for research. Similarly, researchers with limited resources can now conduct high-quality experiments and generate extensive datasets using readily available AI models. The desired form of control for GAI's teacher roles should be behavioural. Referring to the topic modelling example, rather than simply executing the code to generate the topic model, the researcher should inspect and attempt to understand what the GAI has provided. The researcher should query the GAI, if necessary, to seek that understanding. For example, if the code produced is convoluted or opaque, the researcher should request explanations for code segments or request alternate code.

When the GAI is cast in the role of *collaborator*, the human expects it to have equivalent, but different, endowments. Here, the endowments of the GAI and human researcher complement each other. The delegator's objective in invoking the GAI is to seek insight beyond that which the human team members have elicited from the data. This

Table 3. GAI roles in the research process.

GAI role	Relative endowment	Human preference	Performance delegation	GAI control
Assistant	Human > GAI	Efficiency	Human → GAI (command)	Output: human evaluates the end-product generated by GAI
Teacher	GAI > human	Learning	GAI → human (interrogation)	Behavioural: human assesses the process recommended by GAI
Collaborator	Human ≈ GAI	Insight	Human ↔ GAI (iteration)	Clan: human iteratively surfaces ‘rules’ underlying GAI recommendation/output, reconciles those rules with their own, and modifies prompts until rules instantiated by GAI are consistent with those preferred by the human

role of GAI-as-collaborator is exemplified in use of GAI as an open coder in grounded theory research, where the GAI is able to surface codes that humans may not have. Similarly, GAI might complement a researcher’s methodological skills. Such complementarity may be leveraged, for instance, to help researchers apply multi-method research more efficiently and effectively. Also, in the domain of engaged research, GAI models might take the role of co-designers and help researchers to effectively navigate through the solution space for generating useful artefacts.

In these types of engagements with GAI, clan control can be effective. ‘A clan is a culturally homogeneous organisation in which members share common values, beliefs, and norms about how to coordinate effort in order to reach common objectives’ (Sesay et al. 2024: 6). Applying clan control to GAI delegations entails mutual adjustment of values and beliefs. This is necessarily an iterative activity. In it, the researcher first seeks to understand why the GAI offered a particular recommendation by interrogating it repeatedly until the logic for the recommendation is revealed and then restructures prompts to reflect underlying values synthesised from the researcher’s own initial premises and those revealed by their interrogation of the GAI. Just as human collaborators on a grounded theory project iteratively develop and reconcile their codes, drawing upon the norms of grounded theory methods and collaboratively arriving at shared norms for meaningful insight from their data, humans may similarly iterate with a GAI coder to clarify grounded theory norms and collaboratively develop a shared understanding of what constitutes insight about the focal phenomenon. Similarly, researchers undertaking design or engaged science may seek to learn ‘rules of engagement’ through their interactions with GAI and synthesise them with their own explicit or implicit rules into artefacts developed.

Researchers thus must be mindful of the nature of their engagement with GAI – that is, the nature of their tasks and manner in which they are delegating the tasks to the GAI – and implement controls accordingly. Failure to do so can result not only in negative consequences for the research project at hand, but also play to concerns about the possible

long-term detrimental effects of GAI on research, such as researchers losing their ability for critical thinking and to formulate scholarly questions or arguments (Stokel-Walker and van Noorden, 2023).

Finally, the academic enterprise is a community of practice (Wenger et al., 2002), with students entering the enterprise as peripheral participants. When the enterprise works at its best, students apprentice with and learn from mentor experts, progressively moving toward the centre of the community and themselves becoming mentors as they collaborate successfully on the production of scientific knowledge artifacts. GAI now have the potential to enable students at the periphery of the knowledge community to bypass the apprenticeship process in the production of knowledge artifacts. If, in doing so, they assign the GAI the role of teacher and actively seek to learn, their progress in the production of knowledge artifacts will be accompanied by their legitimate progression toward the centre of their knowledge community. If, on the other hand, students leverage GAI as an assistant, they may enjoy serendipitous successes in their production of knowledge artifacts. Those such successes may be rare events, the status afforded to those who enjoy these successes can create a structural disconnect within the community. Specifically, these ‘experts’ – with publication successes – may, in fact, lack the hard-won expertise necessary to mentor the next generation of scholars. Academia needs to plan for such possibilities and develop strategies for ensuring the integrity of our communities.

Concluding remarks

Within the scope of this special issue, we observed a trend towards addressing the challenges of temporal dynamics, complexity, and ethical considerations more effectively in future research methods. We have witnessed a trend towards CTC which is likely to be accelerated with the upcoming developments in the field of GAI. The editors of this special issue believe that these challenges and ethical considerations will continue to be controversial topics. However, it is crucial to balance the promise of GAI with an acute

awareness of the associated risks and maintaining the integrity and authenticity of the research process.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Ivo Blohm  <https://orcid.org/0000-0003-2422-5952>

Shuk Ying Ho  <https://orcid.org/0000-0002-9450-5776>

References

- Abbasi A, Parsons J, Pant G, et al. (2024) Pathways for design research on artificial intelligence. *Information Systems Research* 35(2): 441–459.
- Ataei M, Cheong H, Grandi D, et al. (2025) Elicitron: an LLM agent-based simulation framework for design requirements elicitation. *Journal of Computing and Information Science in Engineering* 25(2): 021012.
- Bail CA (2024) Can Generative AI improve social science? *Proceedings of the National Academy of Sciences of the United States of America* 121(21): e2314021121.
- Baird A and Maruping LM (2021) The next generation of research on IS use: a theoretical framework of delegation to and from agentic IS artifacts. *MIS Quarterly* 45(1): 315–341.
- Banker S, Chatterjee P, Mishra H, et al. (2024) Machine-assisted social psychology hypothesis generation. *American Psychologist* 79(6): 789–797.
- Baskerville RL and Wood-Harper AT (1996) A critical perspective on action research as a method for information systems research. *Journal of Information Technology* 11(3): 235–246.
- Berente N, Seidel S and Safadi H (2019) Data-driven computationally intensive theory development. *Information Systems Research* 30(1): 50–64.
- Berger J (2024) Machines, Psychology, and hypothesis-generation: comment of banker et al. (2024). *American Psychologist* 79(6): 798–799.
- Bilgram V and Laarmann F (2023) Accelerating innovation with generative AI: AI-augmented digital prototyping and innovation methods. *IEEE Engineering Management Review* 51(2): 18–25.
- Bouschery SG, Blazevic V and Piller FT (2023) Augmenting human innovation teams with artificial intelligence: exploring transformer-based language models. *Journal of Product Innovation Management* 40(2): 139–153.
- Burton-Jones A, Butler BS, Scott SV, et al. (2021) Next-generation information systems theorizing: a call to action. *MIS Quarterly* 45(1): 301–314.
- Cardon R and Bibal A (2023) On operations in automatic text simplification. In: Proceedings of the second workshop on text simplification, accessibility and readability, Varna, Bulgaria, 7 September 2023, 116–130.
- Cecez-Kecmanovic D, Davison RM, Fernandez W, et al. (2020) Advancing qualitative IS research methodologies: expanding horizons and seeking new paths. *Journal of the Association for Information Systems* 21(1): 246–263.
- Chen Z and Chan J (2024) Large language model in creative work: the role of collaboration modality and user expertise. *Management Science* 70(12): 9101–9117.
- Chomsky N, Roberts I and Watumull J (2023) *The False Promise of ChatGPT*. New York: New York Times. Available at: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>.
- Chopra F and Haaland I (2023) Conducting qualitative interviews with AI. CESifo Working Paper No. 10666, Center for Economic Studies.
- Creswell JW (2007) *Qualitative inquiry and reserach design: Choosing among five approaches*. 2nd ed. Sage Publications, Inc.
- Croidieu G and Kim PH (2018) Labor of love: amateurs and lay-expertise legitimation in the early US radio field. *Administrative Science Quarterly* 63(1): 1–42.
- Davison RM, Chughtai H, Nielsen P, et al. (2024) The ethics of using generative AI for qualitative data analysis. *Information Systems Journal* 34(5): 1433–1439.
- Doshi AR and Hauser OP (2024) Generative AI enhances individual Creativity but reduces the collective Diversity of novel Content. *Science Advances* 10(28): Eadn5290.
- Durcikova A, Miranda S, Jensen M, et al. (2024) United we stand, divided we fall: an autogenic perspective on empowering cybersecurity in organizations. *MIS Quarterly* 48(4): 1503–1596.
- Fok R, Chang JC, August T, et al. (2024) Qlarify: Recursively expandable abstracts for dynamic information retrieval over scientific papers. In: *Proceedings of the 37th annual ACM symposium on user interface and software*, Pittsburgh, PA, USA, 13-16 Oct 2024.
- Fui-Hoon Nah F, Zheng R, Cai J, et al. (2023) Generative AI and ChatGPT: applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research* 25(3): 277–304.
- George G, Osinga EC, Lavie D, et al. (2016) Big data and data science methods for management research. *Academy of Management Journal* 59(5): 1493–1507.
- Gerow JE, Grover V, Roberts N, et al. (2011) The diffusion of second generation statistical techniques in information systems research from 1990–2008. *Journal of Information Technology Theory and Application* 11(4): 5–28.

- Giunti G and Doherty CP (2024) Cocreating an automated mHealth apps systematic review process with generative AI: design science research approach. *JMIR Medical Education* 10: e48949.
- Goodin D (2024) Hacker plants false memories in ChatGPT to steal user data in perpetuity. *Ars Technica*. <https://arstechnica.com/security/2024/09/false-memories-planted-in-chatgpt-give-hacker-persistent-exfiltration-channel/> last accessed: 1st January 2025.
- Gordetzki P, Blohm I and Hofstetter R (2023) Generative AI in idea development: the role of numeric and visual feedback. In: International conference on information systems, Hyderabad, India, 10–13 Dec 2023.
- Grimes M, Von Krogh G, Feuerriegel S, et al. (2023) From scarcity to abundance: scholars and scholarship in an age of generative artificial intelligence. *Academy of Management Journal* 66(6): 1617–1624.
- Gunning D, Stefik M, Choi J, et al. (2019) Explainable artificial intelligence. *Science Robotics* 4(37): eaay7120.
- Gupta P, Ding B, Guan C, et al. (2024) Generative ai: a systematic review using topic modelling techniques. *Data and Information Management* 8(2): 100066.
- Haman M and Školník M (2023) Using ChatGPT to conduct a literature review. *Accountability in Research* 31(8): 1244–1246.
- Haque E (2024) *AI Horizons: Shaping a Better Future through Responsible Innovation and Human Collaboration*. Sterling, VA: Stylus Publishing, LLC.
- Jain R and Jain A (2024) Generative ai in writing research papers: a new type of algorithmic bias and uncertainty in scholarly work. In: Arai K. (eds). *Intelligent Systems and Applications. IntelliSys 2024. Lecture Notes in Networks and Systems*, Springer Nature, p. 656–669.
- Jarvenpaa S and Klein S (2024) New frontiers in information systems theorizing: human-gAI collaboration. *Journal of the Association for Information Systems* 25(1): 110–121.
- Johnson SL, Gray P and Sarker S (2019) Revisiting IS research practice in the era of big data. *Information and Organization* 29(1): 41–56.
- Kerr NL (1998) HARKing: hypothesizing after the results are known. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc* 2(3): 196–217.
- Khattab O, Singhvi A, Maheshwari P, et al. (2023) DSPy: compiling declarative language model calls into self-improving pipelines. In: Workshop on robustness of zero/few-shot learning in foundation models, *NeurIPS 2023*, New Orleans, LA, 15 Dec 2023.
- Kim HJ, Yang JH, Chang DG, et al. (2024) Assessing the reproducibility of the structured abstracts generated by ChatGPT and bard compared to human-written abstracts in the field of spine surgery: comparative analysis. *Journal of Medical Internet Research* 26: e52001.
- Kulkarni M, Mantere S, Vaara E, et al. (2024) The future of research in an artificial intelligence-driven world. *Journal of Management Inquiry* 33(3): 207–229.
- Lindberg A (2020) Developing theory through integrating human and machine pattern recognition. *Journal of the Association for Information Systems* 21(1): 90–116.
- Locke K, Feldman M and Golden-Biddle K (2022) Coding practices and iterativity: beyond templates for analyzing qualitative data. *Organizational Research Methods* 25(2): 262–284.
- Mansurova M (2023) Topic modelling using ChatGPT API: comprehensive guide to ChatGPT API for newbies. <https://towardsdatascience.com/topic-modelling-using-chatgpt-api-8775b0891d16> last accessed: 1st January 2025.
- Memmert L, Tavanapour N and Bittner E (2023) Learning by doing: educators’ perspective on an illustrative tool for AI-generated scaffolding for students in conceptualizing design science research studies. *Journal of Information Systems Education* 34(3): 279–292.
- Miranda S, Berente N, Seidel S, et al. (2022a) Editor’s comments: computationally intensive theory construction: a primer for authors and reviewers. *MIS Quarterly* 46(2): iii–xviii.
- Miranda S, Wang D and Tian C (2022b) Discursive fields and the diversity-coherence paradox: an ecological perspective on the blockchain community discourse. *MIS Quarterly* 46(3): 1421–1452.
- Mousavi R and Gu B (2024) Resilience messaging: the effect of governors’ social media communications on community compliance during a public health crisis. *Information Systems Research* 35(2): 505–527.
- Nawaz N and Gomes AM (2019) Artificial intelligence chatbots are new recruiters. *International Journal of Advanced Computer Science and Applications* 10(9): 1–5.
- Ngwenyama O and Rowe F (2024) Should we collaborate with AI to conduct literature reviews? Changing epistemic values in a flattening world. *Journal of the Association for Information Systems* 25(1): 122–136.
- Nicholson JM, Mordaunt M, Lopez P, et al. (2021) Scite: a smart citation index that displays the context of citations and classifies their intent using deep learning. *Quantitative Science Studies* 2(3): 882–898.
- Parker JL, Richard VM and Becker K (2023) Guidelines for the integration of large language models in developing and refining interview protocols. *Qualitative Report* 28(12): 3460–3474.
- Pattyn F (2024) The value of generative AI for qualitative research: a pilot study. *Journal of Data Science and Intelligent Systems* (forthcoming).
- Peppers K, Tuunanen T, Rothenberger MA, et al. (2007) A design science research methodology for information systems research. *Journal of Management Information Systems* 24(3): 45–77.
- Perkins M and Roe J (2024) The use of generative AI in qualitative analysis: inductive thematic analysis with ChatGPT. *Journal of Applied Learning and Teaching* 7(1): 390–395.
- Piccoli G, Rodriguez J and Grover V (2022) Digital strategic initiatives and digital resources: construct definition and future research directions. *MIS Quarterly* 46(4): 2289–2316.
- Pratt MG, Kaplan S and Whittington R (2019) Editorial essay: the tumult over transparency: decoupling transparency from

- replication in establishing trustworthy qualitative research. *Administrative Science Quarterly* 65(1): 1–19.
- Rahman MM, Terano HJ, Rahman MN, et al. (2023) ChatGPT and academic research: a review and recommendations based on practical examples. *Journal of Education, Management and Development Studies* 3(1): 1–12.
- Raisch S and Krakowski S (2021) Artificial intelligence and management: the automation–augmentation paradox. *Academy of Management Review* 46(1): 192–210.
- Rossi S, Rossi M, Mukkamala RR, et al. (2024) Augmenting research methods with foundation models and generative AI. *International Journal of Information Management* 77(August 2024): 102749.
- Sabherwal R and Grover V (2024) The societal impacts of generative artificial intelligence: a balanced perspective. *Journal of the Association for Information Systems* 25(1): 13–22.
- Salah M, Abdelfattah F, Alhalbusi H, et al. (2024) Can GAI craft variable Questions? A mixed-method study on AI's capability to adopt, adapt, and create new Scales. https://papers.ssm.com/sol3/papers.cfm?abstract_id=4931424 last accessed: 8th May 2025.
- Sarker S, Xiao X and Beaulieu T (2013) Qualitative studies in information systems: a critical review and some guiding principles. *MIS Quarterly* 37(4): iii–xviii.
- Sarker S, Xiao X, Beaulieu T, et al. (2018) Learning from first-generation qualitative approaches in the IS discipline: an evolutionary view and some implications for authors and evaluators (Part 1/2). *Journal of the Association for Information Systems* 19(8): 752–774.
- Sætre AS and Van de Ven A (2021) Generating theory by abduction. *Academy of Management Review* 46(4): 684–701.
- Saunders C, Van Slyke C and Vogel DR (2004) My time or yours? Managing time visions in global virtual teams. *Academy of Management Perspectives* 18(1): 19–37.
- Schlagwein D and Willcocks L (2023) ChatGPT et al.'s: The ethics of using (generative) artificial intelligence in research and science. *Journal of Information Technology* 38(3): 232–238.
- Sesay A, Ramirez R, Boudreau M-C, et al. (2024) Bureaucratizing the clan: impact of technology affordances on control. *Journal of Management Information Systems* 41(1): 5–38.
- Shmueli G and Koppius OR (2011) Predictive analytics in information systems research. *MIS Quarterly* 35(3): 553–572.
- Shrestha YR, He VF, Puranam P, et al. (2020) Algorithm supported Induction for building theory: how can we use prediction models to theorize? *Organization Science* 32(3): 856–880.
- Shumailov I, Shumaylov Z, Zhao Y, et al. (2023) The curse of recursion: training on generated data makes models forget. arXiv preprint arXiv:2305.17493.
- Sidaoui K, Jaakkola M and Burton J (2020) AI feel you: customer experience assessment via chatbot interviews. *Journal of Service Management* 31(4): 745–766.
- Sinha R, Solola I, Nguyen H, et al. (2024) The role of generative AI in qualitative research: GPT-4's contributions to a grounded theory analysis. In: Proceedings of the symposium on learning, design and technology. Delft, NL, 21 June 2024.
- Smith G (2020) Data mining fool's gold. *Journal of Information Technology* 35(3): 182–194.
- Stokel-Walker C and Van Noorden R (2023) The promise and peril of generative AI. *Nature* 614(1): 214–216.
- Susarla A, Gopal R, Thatcher JB, et al. (2023) The janus effect of generative AI: charting the path for responsible conduct of scholarly activities in information systems. *Information Systems Research* 34(2): 399–408.
- Tingelhoff F, Brugger M and Leimeister JM (2024) A guide for structured literature reviews in business research: the state-of-the-art and how to integrate generative artificial intelligence. *Journal of Information Technology* 40(1): 77–99.
- Tuunanen T, Winter R and Vom Brocke J (2024) Dealing with complexity in design science research: a methodology using design echelons. *MIS Quarterly* 48(2): 427–458.
- Umer F and Adnan N (2024) Generative artificial intelligence: synthetic datasets in dentistry. *BDJ Open* 10(13): 1–5.
- Van Dis EA, Bollen J, Zuidema W, et al. (2023) ChatGPT: five priorities for research. *Nature* 614(7947): 224–226.
- Villalba AC, Brown EM, Scurrill JV, et al. (2023) Automated interviewer or augmented survey? Collecting social data with large language models. arXiv preprint arXiv:2309.10187.
- Volk SC, Schafer MS, Lombardi D, et al. (2024) How generative artificial intelligence portrays science. *Public Understanding of Science* 34(2): 132–153.
- Wenger E, McDermott R and Snyder WM (2002) *Cultivating Communities of Practice: A Guide to Managing Knowledge*. Brighton, MA: Harvard Business Review Press.
- White J, Fu Q, Hays S, et al. (2023) A prompt pattern catalog to enhance prompt engineering with ChatGPT. In: Proceedings of the 30th conference on pattern languages of programs, Monticello, IL, 22-25 Oct 2023.
- Zhang C, Zhang C, Li C, et al. (2023) One small step for generative AI, one giant leap for AGI: a complete Survey on ChatGPT in AIGC Era, arXiv preprint, arXiv preprint arXiv:2304.06488.
- Zheng L, Chiang WL, Sheng Y, et al. (2024) Judging LLM-as-a-judge with MT-bench and chatbot arena. *Advances in Neural Information Processing Systems* 36: 46595–46623.

Author biographies

Ivo Blohm is an Associate Professor for Information Systems and Business Analytics at the Institute of Information Systems and Digital Business at the University of St. Gallen. He studied at Technische Universität München from which he also obtained a doctorate in information systems. His research focuses on leveraging business analytics in organizations to improve decision-making, collaboration, work, and innovation processes. His work has appeared, among others, in such journals as Information Systems Research, Journal of Management Information Systems, Entrepreneurship: Theory and Practice, and Harvard Business Review.

Shaila M. Miranda is M.D. Matthews Endowed Chair and Chair of the Information Systems Department at the University of Arkansas Sam M. Walton College of Business. She has a doctorate in Information Systems (with a minor in Computer Science) from the University of Georgia and an M.A. in Sociology from Columbia University. Her research focuses on the constitutive nature of discourse and on emergent social structures in organizations, digital activism, and digital innovation communities. She employs a combination of qualitative and computational abductive techniques in her research. Shaila's research has appeared in journals such as the *MIS Quarterly*, *Information Systems Research*, and the *Journal of Information Technology*. Her research has been funded by the Department of Homeland Security. She has served as Senior Editor for *MIS Quarterly* and *Information Systems Research* and is a Fellow of the AIS.

Shuk Ying Ho is a Professor at the Australian National University. She received her PhD in Information Systems from the Hong Kong University of Science and Technology. Her research explores human-machine interaction, decision-making processes, and the use of digital trace data for theory development in information systems. She combines

quantitative and computational methods in her work. Her publications have appeared in journals such as *MIS Quarterly*, *Information Systems Research*, and the *Journal of the Association for Information Systems*. She serves as a Senior Editor for *MIS Quarterly* and *AIS Transactions on Human-Computer Interaction*, and sits on the editorial boards of the *Journal of the Association for Information Systems* and the *Journal of Management Information System*.

Jan Marco Leimeister is Managing Director and Chair Professor at the Institute of Information Systems and Digital Business, University of St. Gallen, Switzerland. He is furthermore Professor and Director of the Research Center for Information System Design (ITeG) at the University of Kassel, Germany. His research covers Digital and AI driven Business, Digital Transformation, Service Innovation and Management, Future of Digital Work, Teaching and Learning and Hybrid Intelligence. He runs research groups and projects funded by the European Union, national ministries, DFG, SNF, various foundations, and industry partners. He is a co-founder of several companies and serves as board member, consultant, coach, and speaker for numerous organizations.