Rieskamp, Jonas; Budnik, Yuliya; Mirbabaie, Milad

Why AI Deployment Fails in Organisations : A Socio- Technical Perspective on the Root Causes

Date of secondary publication: 04.05.2026

Version of Record (Published Version), Conferenceobject

Persistent identifier: urn:nbn:de:bvb:473-irb-114927x

Primary publication

Rieskamp, Jonas; Budnik, Yuliya; Mirbabaie, Milad (2025): Why AI Deployment Fails in Organisations: A Socio-Technical Perspective on the Root Causes, in: ACIS 2025 Proceedings, AIS

Why AI Deployment Fails in Organisations: A Socio-Technical Perspective on the Root Causes

Full research paper

Jonas Rieskamp
Faculty of Information Systems and Applied Computer Sciences
University of Bamberg, Germany
Bamberg, Germany
Email: jonas.rieskamp@uni-bamberg.de

Yuliya Budnik
Faculty of Information Systems and Applied Computer Sciences
University of Bamberg, Germany
Bamberg, Germany
Email: yuliya.budnik@stud.uni-bamberg.de

Milad Mirbabaie
Faculty of Information Systems and Applied Computer Sciences
University of Bamberg, Germany
Bamberg, Germany
Email: milad.mirbabaie@uni-bamberg.de

Abstract

AI deployment challenges often arise from misalignments between technological design and the social systems into which it is introduced. Drawing on Bostrom and Heinen's (1977a) Socio-Technical Systems framework, this study analyses contemporary AI literature to identify the root causes of recurring failures. The work adapts the framework into a diagnostic tool that links socio-technical conditions to common AI challenges, offering both a conceptual lens for research and practical guidance for governance and design. While literature-based, the framework provides a foundation for empirical testing to improve proactive, context-sensitive AI integration.

Keywords Artificial Intelligence, Socio-Technical Systems, Diagnostic Framework, AI Governance, Organisational Change Management

1 Introduction

The dominant narrative surrounding artificial intelligence (AI) in organisations is one of unprecedented efficiency and transformative potential (Collins et al. 2021). By automating routine tasks, AI promises to free employees for more complex, creative, and value-generating work (Makarius et al. 2020; Raisch & Fomina 2025). This widespread optimism is reflected in staggering global investments exceeding $140 billion annually, with projections estimating a potential economic impact of up to $15 trillion by 2030 (Jeon et al, 2024). Yet, alongside these promises, AI adoption in organisations also reshapes existing power dynamics (Hoddinghaus et al. 2021; Qin et al. 2023). It raises questions about who gains access to the efficiency and innovation it promises, whose perspectives shape system design, and how such shifts may reinforce or challenge existing hierarchies.

AI systems don't arrive in a vacuum. They come embedded in expectations about what good work looks like, what productivity means, and what kind of behaviour can—or should—be automated. While these are fundamentally mere tools in theory, in practice, they frequently function as silent architects shaping both organisational structures and social dynamics. Many now lie at the heart of what we call management information systems (MIS), allocating resources (Duda et al. 2024), monitoring performance (Bednar & Welch 2020; Siering 2022), and making recommendations for decisions (Di Noia et al. 2022; Gupta et al. 2022). This point often marks the onset of challenges, as such systems frequently overlook human heterogeneity, thereby perpetuating existing inequalities and giving rise to new forms of exclusion.

Such dynamics create a high-stakes paradox for organisations. On the one hand, immense pressure exists to adopt AI to remain competitive, spurred by warnings that "managers who use AI will replace those who don't" (Brynjolfsson and McAfee (2017), as cited in Gupta et al. 2022, p. 1476). On the other hand, the widespread neglect of these socio-technical complexities results in significant implementation problems: project failure (Gyldenkaerne et al. 2024; Weber et al. 2023), employee dissatisfaction (Bednar & Welch, 2020; Thomas, 2024), overreliance on tools, and misplaced trust in algorithmic outputs (Fleiss et al. 2024; Floridi et al. 2018; Klingbeil et al. 2024). Considering recent calls for the deregulation of AI in the United States, it is now more crucial than ever to implement proactive measures to ensure the responsible development and deployment of such technologies (Sahu & Mohanty, 2024).

As early as 1977, Bostrom and Heinen recognised that such failures often result not from inadequate technology but from a fundamental misalignment of design assumptions with behavioural and structural realities (Bostrom & Heinen 1977a; 1977b). Their socio-technical systems (STS) framework offers a powerful diagnostic lens for understanding why even the most efficient innovations can backfire when the social subsystem is overlooked. We posit that their insights are more relevant than ever for navigating the challenges of AI, as many of its most pressing issues typically surface only post-deployment, when corrective measures are most difficult. The central research question guiding this study is therefore:

RQ: How can the STS framework be effectively applied to social and behavioural challenges in AI deployment?

To address this question, this paper employs a systematic literature review to identify and categorise the primary challenges associated with AI. These barriers are then mapped onto the seven socio-technical conditions for failure proposed in the original STS framework, making their root causes explicit. This study's contribution is twofold: First, it validates the enduring relevance of a classic theoretical framework for diagnosing contemporary AI challenges. Second, systematically unearthing the root causes of these challenges provides a crucial diagnostic foundation that must precede the development of any truly effective ethical and managerial interventions.

2 Theoretical Background

MIS projects have a long and well-documented history of failure. These failures stem more often from human and organisational dynamics than from technical shortcomings. As early as 1975, Henry C. Lucas observed that "the major reason most information systems have failed is that we have ignored organisational behaviour problems in the design and operation of computer-based information systems" (Lucas (1975), as cited in Bostrom & Heinen 1977a, p. 18). It was in direct response to his oversight that Bostrom and Heinen (1977) introduced their STS framework. The framework bridges technological design and its social consequences, anchored in a principle of *joint optimisation*: improving technical efficiency and enhancing the quality of working life in tandem.

The diagnostic strength of the STS framework lies in three deceptively simple but powerful questions: (Q1) *What are the human or behavioural challenges?* (Q2) *What are the root causes of these challenges?* and (Q3) *How can these causes be addressed?* Bostrom and Heinen argued that the second question is chronically neglected. Too often, practitioners leap from merely cataloguing symptoms (a “story-telling phase”) straight into proposing ad hoc remedies (the “instant solution phase”), without ever developing a grounded understanding of underlying causes.

To address this blind spot, they identified seven recurring conditions, entrenched mindsets and practices among MIS designers, that give rise to inadequate designs and, ultimately, system failure. The development of concrete prescriptive design methodologies, which they explore in a companion paper (Bostrom & Heinen 1977b), is a separate endeavour that logically follows the foundational analysis. These seven conditions, each reflecting a breakdown in the joint optimisation of social and technical subsystems, form the analytical lens for this study and are summarised in the appendix as a reminder (see Table 2).

2.1 Related Work and Research Problem

The discourse surrounding AI is situated at the intersection of several mature research streams. The STS perspective has long been a valuable perspective for how social and technical elements shape and constrain one another (Trist & Bamforth 1951). Contemporary STS research has evolved to address the complexities of modern digital work, introducing concepts like socio-technical capital in human-AI collaboration (Makarius et al. 2020), the need for socio-technical envelopment to manage opaque AI models (Asatiani et al. 2021), and the challenge that dynamic AI systems pose to traditional notions of organisational stability (Fischer & Baskerville 2023).

Simultaneously, research engages with the organisational realities in AI deployment. Here, success hinges on aligning AI capabilities with strategic objectives, underpinned by a robust, data-driven culture (Fosso Wamba et al. 2024; Sestino & De Mauro 2022). Researchers emphasise the need for disciplined lifecycle management, transparent communication, and sensemaking processes that help employees integrate AI systems into their workflows (Engstrom et al. 2024; Weber et al. 2023). This is complemented by calls for comprehensive AI governance that spans data governance (quality, privacy), model governance (fairness, explainability), and system governance (compliance, performance) (Schneider et al. 2023). Such structures are not purely technical safeguards since they also influence employee perceptions of transparency, fairness, and job security, which in turn shape adoption outcomes (Chiu et al. 2021).

A third, increasingly prominent body of work interrogated the ethical dimensions of AI, often framed through the principles of fairness, accountability, and transparency (FAT). This stream addresses the risks of algorithmic bias, where models perpetuate inequities embedded in historical data (Feuerriegel et al. 2020); the opacity of deep learning models, which impedes interpretability (Ashok et al. 2022); and the challenge of accountability when AI autonomy complicates the attribution of responsibility (Floridi et al. 2018). Privacy concerns, arising from AI’s reliance on large-scale and often sensitive datasets, are a recurring theme (Merhi 2023).

While each of these streams offers valuable insights, together they reveal a pattern that Bostrom and Heinen (1977a; 1977b) famously critiqued. Much of the literature remains anchored either in cataloguing social and behavioural challenges (Q1) or prescribing reactive interventions (Q3), while paying insufficient attention to the second and arguably most crucial question (Q2): *Why do these problems arise in the first place?* Employee resistance, fairness concerns, and other barriers are thoroughly described, and governance mechanisms are extensively proposed; however, the underlying drivers, such as flawed design assumptions, unintended systemic effects, or misaligned organisational incentives, often receive only cursory treatment.

Our study seeks to approach that problem. By systematically applying Bostrom and Heinen’s (1977) STS framework, it examines the socio-technical challenges of AI through the lens of their root causes and maps them against the framework’s seven conditions for failure. The goal here is not yet to prescribe definite solutions for every barrier but to surface the deeper, structural factors that shape implementation outcomes. In doing so, we aim to enrich understanding and lay a more solid foundation for interventions that are both technically sound and socially attuned.

3 Research Approach

This study employs a systematic literature review based on the concept-driven methodology established by Webster and Watson (2002). This approach was chosen to systematically synthesise the fragmented, multidisciplinary literature on AI's organisational and social challenges. The entire process follows the PRISMA standard (Page et al. 2021) to ensure a transparent and replicable research design.

3.1 Literature Search and Selection

The literature search was conducted in the Web of Science (WoS) and AIS Electronic Library (AISeL) databases, with the scope restricted to a predefined list of 51 core information systems and related journals (Boell & Wang 2019). The search query combined keywords for artificial intelligence with terms related to social and ethical issues (see Table 3). The selection process followed a multi-stage protocol (see Figure 2): after an initial screening of titles and abstracts, full texts were assessed against specific inclusion criteria. We included empirical or substantive theoretical studies addressing the behavioural and social dimensions of AI, while excluding purely technical papers. This process, supplemented by citation tracking, resulted in a final corpus of 91 studies (see Table 4).

3.2 Data Analysis and Conceptual Synthesis

We employed a concept-centric coding process, adapted from Wolfswinkel et al. (2013), which combines deductive and inductive phases. The analysis began with a set of deductive categories derived from foundational STS literature (e.g., bias, digital divide, trust). Subsequently, each study was analysed via open coding to inductively identify emergent concepts from the data. This allowed for the refinement of abstract notions (e.g., "fairness") into more concrete analytical concepts (e.g., "bias", "algorithmic aversion").

In the final synthesis, these concepts were consolidated through a multi-level clustering process. First, thematically related codes were grouped into 17 distinct sub-clusters. To ensure the robustness of our findings, only concepts identified in three or more independent sources were included. These sub-clusters were then aggregated into four overarching analytical clusters that structure our results: (1) Reproduction of Structural Inequality, (2) Erosion of Trust, (3) Erosion of Meaningful Work, and (4) Consolidation of Power (see Figure 1 and Table 3). Together, these clusters illustrate how AI systems, when misaligned with the organisational and social contexts they inhabit, produce significant social and behavioural consequences.

4 Results

4.1 Structural Inequality

AI systems often do not merely reflect existing inequality but amplify it. The most prominent mechanism is algorithmic *bias*, which is not a simple coding error, but a context-dependent phenomenon deeply rooted in societal structures (Buyl & de Bie 2024; Feuerriegel et al. 2020; Koechling et al. 2021). Bias can become self-perpetuating: feedback loops generate increasingly skewed datasets, entrenching structural discrimination (Akter, McCarthy, et al. 2021; Bauer et al. 2024; Bedemariam & Wessel 2023; Gupta et al. 2022; Hoffmann 2019). The risk of fairness gerrymandering exemplifies this dynamic, whereby aggregate fairness targets are met while specific subgroups (often those at the intersection of multiple marginalised identities) remain systematically disadvantaged (Buyl & de Bie 2024).

These mechanisms have tangible, identity-specific effects. In gendered contexts, talent acquisition systems frequently reinforce occupational hierarchies by favouring career paths dominated by men (Aydin et al. 2023; Hoffmann 2019; Yarger et al. 2020). Behavioural patterns differ as well: Women are more likely to scrutinise biased algorithmic decisions, even when accompanied by transparent explanations (Gupta et al. 2022; Tsung-Yu et al. 2024; van Berkel et al. 2023). Binary classification schemes also systematically exclude or misgender transgender and non-binary individuals (Keyes 2018). Ethnicity becomes another vector of inequity, particularly in hiring (Koechling et al., 2021; Kordzadeh & Ghasemaghaei 2022; Yarger et al. 2020), healthcare (Birzhandi & Cho 2023; Isaac et al. 2024; Moussawi et al. 2024), and criminal justice (Birzhandi & Cho 2023; Buyl & de Bie 2024; Feuerriegel et al. 2020; Gupta et al. 2022). For example, some minority applicants adopt counterstrategies such as *code-switching* to navigate systems they perceive as biased (Bedemariam & Wessel 2023). As a result, repeated experiences of automated rejection foster heightened distrust, reduced engagement, and increased feelings of procedural unfairness (Bedemariam & Wessel 2023; Bigman et al. 2021; Yarger et al. 2020).

Inequality extends to age and disability. Systems designed with younger, tech-savvy users in mind can alienate older adults, fostering scepticism and lower adoption rates (Birzhandi & Cho 2023; Fleiss et al. 2024; Siering 2022). Similarly, AI optimised for able-bodied norms often fails to accommodate people with disabilities, and “fairness through awareness” (excluding disability-related data) proves ineffective

(Buyl & de Bie 2024; Feuerriegel et al. 2020) if proxy variables such as speech patterns continue to disadvantage these users (Birzhandi & Cho 2023).

The digital divide further exacerbates structural exclusion. Limited access to digital literacy, infrastructure, and resources prevents marginalised groups from leveraging AI benefits, while resource-rich entities expand their competitive lead (Carter et al. 2020; Heng et al. 2022; Jeon et al. 2024; Marjanovic et al. 2022; Ransbotham et al. 2016). The result is a widening socio-economic gap that risks becoming permanently entrenched.

4.2 Erosion of Trust

A second cluster centres on the social and psychological processes that erode user trust in AI. Foremost among these is the perceived *breach of privacy*, driven by limited transparency and user control over personal data (Kushwaha et al. 2023; Sipior et al. 2024; Zhdanov et al. 2022). Personalisation can heighten this tension: While it offers some value, it requires extensive data sharing, which in turn deepens privacy concerns (Buyl & de Bie 2024; Liu & Tao 2022). The tension between privacy and system performance becomes especially clear in data-driven business models. Here, privacy measures such as data anonymisation can inadvertently impair the predictive accuracy of AI models by restricting access to crucial training data (Laine et al. 2024; Sipior et al. 2024; Zhdanov et al. 2022).

Trust is further undermined by *algorithmic aversion*, where users perceive algorithmic decisions as less fair than human ones, even when more accurate (Bankins et al. 2022; Hoddinghaus et al. 2021; Klingbeil et al. 2024; Langer & Landers 2021; Qin et al. 2023). The „transparency-resistance paradox" furthermore captures that excessive transparency in complex contexts can overwhelm rather than reassure users, ultimately reducing their trust (Hu et al. 2024). Aversion intensifies when AI is seen as threatening professional identity or competence (Mirbabaie et al. 2022; Qin et al. 2023). Female participants demonstrated greater awareness of discrimination and less trust in biased AI systems compared to male participants, especially when information-rich explanations were provided (Tsung-Yu et al. 2024). Additionally, older users tend to have less trust in algorithmic decisions (Gude 2023; D. Kim & Song 2024). Perceptions of AI discrimination, trust, and fairness vary systematically: liberals are more sensitive yet more accepting of AI, conservatives are more sceptical yet less emotionally reactive to discriminatory outcomes; digital self-efficacy boosts confidence, not fairness; and higher socioeconomic status fosters an optimistic bias that downplays algorithmic risks (S. Kim 2025).

Anthropomorphic design choices can also backfire. While human-like cues may initially engage users, inconsistencies or inauthentic interactions can trigger an *uncanny valley* effect, leading to discomfort and distrust (Hoorn & Huang 2024). Conversely, some users form parasocial relationships with AI agents, accepting their recommendations less critically due to perceived emotional bonds (Noor et al. 2022; M. Sharma et al. 2024). Cultural background moderates these effects: Users from individualistic cultures often respond more positively to anthropomorphism than those from collectivist backgrounds (Chang et al. 2018). Notably, anthropomorphic chatbots promote greater personal information disclosure (Schanke et al. 2021). In governance and organisational decision-making, anthropomorphic AI often reinforces existing social hierarchies and biases (Oleksy et al. 2023). For instance, female-voiced virtual assistants and other gendered designs have been proven to strengthen gender stereotypes and contribute to professional inequalities (Mou & Xu 2017; Tsung-Yu et al. 2024).

Ultimately, trust erosion often culminates in *dehumanisation*. When algorithms reduce individuals to data points and ignore personal context, users report emotional detachment and diminished social value (Bankins et al. 2022; Bedemariam & Wessel 2023; Langer & Landers 2021). In professional and social environments, this risk replaces genuine connection with mechanised, emotionally hollow exchanges (Porra et al. 2020).

4.3 Erosion of Meaningful Work

The degradation of meaning, value, and security in human labour is mostly visible in *job displacement* (Loureiro et al. 2021; Y. Lu 2019; Marabelli et al. 2021), now extending beyond routine work into roles requiring education, expertise (Zhang et al., 2024), and managerial discretion (Mirbabaie et al. 2025).

Further, *overreliance* (i.e., excessive trust in automated outputs) diminishes critical engagement with AI's output (Ghasemaghaei & Kordzadeh 2024a; 2024b; Klingbeil et al. 2024; Qin et al. 2023). This can result in deskilling, as domain knowledge and judgment atrophy (Floridi et al. 2018; Langer & Landers 2021; Mirbabaie et al. 2022). Explainability, while often seen as a means for critical interrogation, can also encourage uncritical acceptance of biased results when presented in a plausible manner (Ghasemaghaei & Kordzadeh 2024a). In one study, individuals with lower fairness sensitivity perceived the AI as fairer solely due to its transparency, even when it explicitly disclosed gender bias (Tsung-Yu et al., 2024). The perception of scientific legitimacy can reduce decision-makers' willingness to consider alternative solutions or intervene when AI outputs are flawed (Cowgill et al. 2020). This behaviour leads to decreased engagement in decision processes (Floridi et al. 2018; Klingbeil et al. 2024; Loureiro et al. 2021), resulting in errors of omission—failing to step in when AI overlooks important issues—and errors of commission—wrongly accepting incorrect algorithmic recommendations (Sipior et al. 2024).

These dynamics frequently lead to a *loss of control*. Employees feel their autonomy eroded when AI systems override professional judgment without recourse (Bankins et al. 2022; Bedemariam & Wessel 2023; Floridi et al. 2018; Langer & Landers 2021; Y. Lu 2019; Marjanovic et al. 2022; Marotta 2022; Wesche & Sonderegger 2021). Compounding this is a dilution of accountability: As responsibility is diffused across complex AI systems, it becomes unclear who answers for errors (Floridi et al. 2018; Langer & Landers 2021; Marjanovic et al. 2022; Marotta 2022). Feelings of limited influence over outcomes often spark resistance toward AI systems (Mirbabaie et al. 2022), particularly when unexpected or opaque decisions force users to speculate about or anticipate the system's behaviour, in workplace settings, this dynamic reinforces a feedback loop in which human expertise is increasingly undervalued as artificial agents take over more complex decision-making roles (Floridi et al. 2018; Sartori & Theodorou 2022).

4.4 Consolidation of Power

AI consolidates and centralises power at multiple levels. Globally, this is visible in the emerging concept of *data colonialism*, where data is extracted from the Global South, but economic and innovative value accrues primarily to the Global North's tech hubs (Arora et al. 2023; Heeks et al. 2024; Kwet 2019; Seoane & Velasco 2024). This process not only entrenches economic asymmetries but also extends cultural influence as Western-trained models diffuse their embedded norms globally, reproducing hegemonic world views (Rieskamp et al. 2023; van Berkel et al. 2023).

Within organisations, AI shifts *power dynamics* by granting management unparalleled oversight and decision-making leverage under the pretence of AI's objectivity (Floridi et al. 2018; Someh et al. 2019). Promoted by remote work trends, algorithmic monitoring can determine career trajectories and compensation without transparent human input, fostering disempowerment (Marabelli et al., 2021; Sipior et al. 2024; van Berkel et al. 2023). The complexity of AI systems limits employee oversight, contributing to an environment where algorithmic decisions are perceived as inevitable and unchallengeable (van Berkel et al. 2023). Relying on AI to manage performance and outcomes strengthens managerial control, as it diminishes employee bargaining power (Floridi et al. 2018; Marabelli et al. 2021).

Central to this is *algorithmic control*: Scalable, often invisible forms of surveillance and profiling (Floridi et al. 2018; Heeks et al. 2024; Seoane & Velasco 2024; Someh et al. 2019). Surveillance involves active monitoring through tools like facial recognition, while profiling individuals for critical decisions in law enforcement, and beyond (Buyl & de Bie 2024; Floridi et al. 2018; Marjanovic et al. 2022; Ransbotham et al. 2016). Excessive data mining is increasingly framed as a form of surveillance, as illustrated by the claim that "Big Data is little more than a euphemism for surveillance" (Kwet 2019, p. 7). Predictive policing illustrates the risks, targeting the same communities repeatedly and using algorithmic "objectivity" to reinforce entrenched biases (Feuerriegel et al. 2020; Floridi et al. 2018; Marjanovic et al. 2022). Exporting AI surveillance systems to low- and middle-income countries creates lasting technological dependence–external providers frequently operate such systems, which diminishes local autonomy and regulatory oversight (Heeks et al. 2024).

5 Discussion

The findings portray the challenges posed by AI as daunting. The sheer number of issues arising from its use can cause unease, echoing deeper societal anxieties that have accompanied previous technological shifts, like the widespread disruption during the Industrial Revolution. However, this dark part of history offers more than just cautionary tales; it provides reasons for cautious optimism as new institutions can emerge to steer innovation toward a more balanced path.

Acknowledging the scale of these challenges is the essential first step. However, to act effectively, we must go beyond simply cataloguing symptoms (Q1). Instead, we need to explore their root causes (Q2). This is the significant task this discussion aims to undertake: By systematically mapping the identified challenges onto the seven conditions outlined by Bostrom and Heinen (1977a) (see Table 1), we seek to reveal the root causes underlying AI deployment failures and thus provide a robust foundation for meaningful action.

Condition	Core Causal Mechanisms (The "How")	Resulting Challenges (The "What")	Solution Principles
C1: Systems Designer's Implicit Theories	Unchallenged assumptions about 'default users' and technological value-neutrality become embedded in system logic	Bias, Dehumanisation	Uncover and challenge implicit theories about users, work, and technology at the earliest design stages.
C2: Systems Designer's Concept of Responsibility	Unclear accountability structures and delegation of moral agency to algorithms rather than humans	Algorithmic Aversion, Algorithmic Control, Loss of Control, Overreliance, Sense of Accountability	Create accountability-by-design and clear override pathways.
C3: Non-Systemic View	Isolated technical optimisation that ignores feedback loops, secondary effects, and broader organisational and societal consequences	Algorithmic Control, Bias, Breach of Privacy, Dehumanisation, Digital Divide, Global Power Dynamics, Job Displacement, Organisational Power Dynamics	Systemically scout for secondary effects before optimising local metrics.
C4: Limited Goal Orientation	A narrow focus on efficiency and cost metrics sidelines social objectives	Bias, Global Power Dynamics, Job Displacement, Overreliance	Implement dual goal-setting that treats efficiency and quality of working life as independent, co-equal targets.
C5: Limited Design Referent Group	Exclusion of marginalised perspectives and secondary users from participatory design processes	Algorithmic Aversion, Bias, Dehumanisation, Digital Divide, Uncanny Valley	Broaden the referent groups and stakeholder ownership deliberately.
C6: Rational/Static View of Development	Assumption of stable context prevents adaptation to concept drift and evolving user behaviours over time	Algorithmic Aversion, Algorithmic Control, Bias, Breach of Privacy, Global Power Dynamics, Loss of Control, Organisational Power Dynamics, Overreliance	Anticipate concept drift and mission creep with iterative, learning-oriented lifecycles.
C7: Limited Change Technologies	Technical deployment without adequate training, role clarification, or integration with organisational workflows and culture	Algorithmic Aversion, Breach of Privacy, Dehumanisation, Digital Divide, Job Displacement, Loss of Control, Overreliance, Sense of Accountability, Uncanny Valley	Integrate a change of technologies from organisational development beyond training and UI tweaks

Table 1 Mapping of Socio-Technical Conditions and the Associated Challenges

5.1 Condition 1: Systems Designer's Implicit Theories

The first condition posits that designers' unspoken assumptions and mental models are a primary source of system failure (Bostrom & Heinen 1977a). Our findings confirm this is a critical origin point for systemic bias in AI systems, as designers' unchallenged beliefs become embedded in the system's logic (Aydin et al. 2023; Gupta et al. 2022; Yarger et al. 2020).

This mechanism manifests directly as *bias* when designers implicitly assume a "default user" and consequently fail to account for those who don't fit this mould. Flawed mental models of fairness itself compound the problem; the attempt to achieve "fairness through unawareness" by omitting sensitive attributes often backfires, since algorithms leverage proxy variables that perpetuate the very discrimination intended to be removed (Birzhandi & Cho 2023; Cornacchia et al. 2023; Feuerriegel et al. 2020). This reveals a deeper, implicit theory: The persistent yet flawed belief that technology is value-neutral, and that bias is merely a technical glitch, rather than a reflection of embedded institutional forces (Marjanovic et al. 2022).

The same condition also fosters *dehumanisation*. When designers hold a view of users as predictable and purely rational actors (or worse, as mere data points in an optimisation problem), they design systems that treat humans in a mathematical, mechanical way (Bankins et al. 2022; Fleiss et al. 2024; Qin et al., 2023). When such AI-driven processes ignore essential personal contexts, users report feeling reduced and invalidated (Bankins et al. 2022; Ochmann et al. 2024), a direct consequence of a design model that prioritises technical optimisation over empathetic, relational processes.

Ultimately, the implicit theories of Condition 1 are not just one cause among seven. Instead, they are often the foundational seed from which many other socio-technical failures grow.

5.2 Condition 2: Systems Designer's Concept of Responsibility

The second condition examines how designers conceptualise accountability within an STS (Bostrom & Heinen, 1977a). Failures arise when designers create systems that obscure human responsibility or implicitly encourage its delegation to the machine. Our findings show that a flawed concept of responsibility is the root cause of two seemingly contradictory yet deeply related barriers: *overreliance* and *algorithmic aversion*.

Overreliance emerges when designers build systems that are perceived as autonomous decision-makers, implicitly framing human users as passive actors. This encourages users to offload responsibility, particularly when algorithmic outputs are presented as objective or authoritative (Ghasemaghaei & Kordzadeh 2024b; Wu et al. 2024). The automation bias exacerbates this dynamic when users excessively trust algorithmic recommendations and uncritically accept flawed or biased outcomes (Fleiss et al. 2024; Sipior et al. 2024). The result is a dangerous erosion of accountability. Managers have been found to follow unjust algorithmic suggestions without guilt, effectively blaming the system for discriminatory decisions (Ghasemaghaei & Kordzadeh 2024b).

Conversely, a lack of clear accountability structures can also lead to *algorithmic aversion*. Savvy users may distrust and resist systems precisely because they recognise that an algorithm cannot be held accountable for its actions (Fleiss et al. 2024; Ransbotham et al. 2016). This effect is tightly linked to a perceived *loss of control*: Employees feel they lack the agency to intervene in or to override automated processes, leaving them powerless with a system that makes decisions without being accountable (Floridi et al. 2018; Mirbabaie et al. 2022).

Both overreliance and aversion, therefore, are symptoms of a central failure: The system's design and implementation did not efficiently establish a structure of shared, human-centred responsibility.

5.3 Condition 3: Non-Systemic View

The third condition highlights a failure to adopt a holistic perspective, where designers optimise a system in isolation, ignoring its interactions and secondary effects within the broader social context (Bostrom & Heinen 1977a). This narrow focus is particularly dangerous in AI deployment because it neglects how AI not only operates within a social framework but also actively shapes it, often through self-reinforcing feedback loops.

The amplification of *bias* is most illustrative of this non-systematic view. An AI model used for predictive policing, trained on historically biased arrest data, will inevitably focus patrols on already over-policed communities if we overlook the origins of the data. This intervention alters social reality: Increased patrols lead to more arrests, which in turn generate more biased data that feeds back into the model (Bauer et al. 2024; Feuerriegel et al. 2020; Gupta et al. 2022), potentially leading to the collapse of AI models (Savage 2025). This creates a self-perpetuating cycle of discrimination that fails to recognise the bias in the data.

The failure to recognise the larger surrounding system has profound systemic consequences across society. On a macro level, it fosters forms of *global power dynamics through data colonialism*, where local data fuels innovation for global tech giants while reinforcing geopolitical dependencies (Arora et al. 2023; Gravett 2022). It also widens the *digital divide* by deploying services that systematically exclude those lacking access or skills (Carter et al. 2020; Jeon et al. 2024). On an organisational level, a narrow focus on efficiency can lead to unintended shifts in *organisational power dynamics* and the implementation of *algorithmic control*, undermining employee autonomy and provoking resistance

(Claure et al. 2023; Heeks et al. 2024; Loureiro et al. 2021). The societal costs of *job displacement* (Floridi et al. 2018; Ransbotham et al. 2016) and individual experiences of *dehumanisation* through efficient but emotionally hollow and automated processes (Porra et al. 2020) trace back to this same fundamental failure to anticipate social ripple effects.

These tensions also illustrate a common challenge in system optimisation, where maximising one metric can inadvertently harm another (e.g., the pursuit of efficiency may compromise equitable outcomes). Similarly, improving fairness through better representation could conflict with *privacy*, as it might require vulnerable groups to provide more sensitive data (Buyl & de Bie 2024; Liu & Tao 2022). Another trade-off exists between trust and transparency. While greater transparency can enhance understanding, it can also overwhelm users, ultimately undermining their trust (Hu et al., 2024). These challenges reveal how optimising for a single goal can create new barriers, emphasising the necessity for a systemic perspective that anticipates such effects.

In sum, Condition 3 failures are responsible for many second-order effects of AI deployment. If designers or organisations do not account for the whole picture, then the technology interacts with existing structures in harmful ways (Langer & Landers 2021; Loureiro et al. 2021). A non-systemic approach favours those who are already in a powerful position. C3 misalignment becomes visible wherever AI implementation produces broad social side effects that designers should have accounted for but didn't.

5.4 Condition 4: Limited Goal Orientation

The fourth condition focuses on a narrow definition of success, in which system implementation is primarily guided by technological and economic goals, such as productivity and cost reduction. This approach often overlooks important social objectives (Bostrom & Heinen 1977a). This limited focus on goals is a major factor contributing to socio-technical failures, since it frequently requires a trade-off between measurable system efficiency and social values.

This tension is particularly evident in the issue of *bias*. Organisations often prioritise predictive accuracy or efficiency, even if it means perpetuating unfair practices (Akter, Dwivedi, et al. 2021; Ghasemaghaei & Kordzadeh 2024a; Koechling et al. 2021; Kordzadeh & Ghasemaghaei 2022; Moussawi et al. 2024; Ransbotham et al. 2016). For example, in algorithmic hiring, an exclusive focus on matching candidates to historical profiles of success can result in the systematic exclusion of non-traditional candidates, thus reinforcing homogeneity in the workplace (Koechling et al. 2021; Yarger et al. 2020). This is frequently not an oversight but rather a strategic choice, driven by institutional pressures to conform to industry norms that prioritise easily quantifiable performance metrics over less tangible social goals. However, this strategy is shortsighted, as it sacrifices the long-term advantages of workforce diversity for the sake of short-term uniformity (Dimaggio & Powell 1983).

Additionally, the pressure to meet performance targets can lead to *overreliance*, where employees may feel compelled to rely on algorithmic recommendations to save time, especially in high-stakes situations or tight deadlines (Klingbeil et al. 2024; Kushwaha et al., 2023; Langer & Landers, 2021). When cost reduction is framed as the ultimate goal, AI implementation may become focused on replacing human workers, leading to *job displacement* (Ransbotham et al. 2016). If organisations do not invest in retraining displaced workers because that's outside the immediate cost metric, employees feel expendable, raising public unemployment issues (related to C3's systemic problems). On a global scale, if the goal is framed as connecting a certain number of users to a service, a company may neglect the simultaneous goal of ensuring that those million users have local data rights and benefits. Hence, global expansion often follows the path of least resistance, which can lead to persistent *global power imbalances* (Arora et al. 2023).

If an entire industry prioritises efficiency above all else, each organisation will align its goals accordingly. This reveals a systemic issue that extends beyond individual designs. It is important to recognise that addressing these issues may require reframing success criteria at the market or societal level. Condition 4 highlights that if fairness and human factors are not explicitly targeted, the system will not achieve those outcomes and render additional systemic shortsightness. A limited goal orientation can also reinforce a non-systemic view (C3) if organisations measure success over narrow metrics instead of accounting for broader systemic consequences. In this sense, C4 often acts as a precursor or driver of C3, as limited goals constrain the scope of what is considered relevant during system optimisation.

5.5 Condition 5: Limited Design Referent Group

Fifth, failures occur when systems are designed exclusively for a narrow group of users, neglecting other stakeholders' needs (Bostrom & Heinen 1977a). If marginalised communities are excluded from the design process, the resulting systems are often ill-suited for them, perpetuating bias (Gupta et al. 2022; Nedungadi et al. 2024). For instance, a recruiting tool trained on biased historical data might compel minority job applicants to engage in "code-switching" or mask their identities to adapt to a system that was not designed with them in mind (Bedemariam & Wessel 2023).

This exclusion breeds mistrust and alienation, creating barriers such as the *uncanny valley* effect, *dehumanisation*, and *algorithmic aversion*. When designers fail to consider the diverse cultural and individual expectations, anthropomorphic systems can feel unsettling or emotionally hollow (Baudier & de Boissieu 2025; Bigman et al. 2021; Chang et al. 2018). Users who realise they were not consulted often perceive the technology as being imposed upon them, leading to suspicion and resistance. Moreover, designing with an idealised, tech-savvy user in mind, while ignoring those with fewer resources or lower proficiency, actively widens the digital divide (Hu et al. 2024).

Notably, Condition 5 pertains to absent perspectives, whereas Condition 1 addresses biased perspectives that are present. These conditions frequently coincide. For example, a homogeneous design team is likely to produce a limited referent group. On the other hand, broadening the referent group (C5) can help designers gain a more comprehensive view and challenge their implicit assumptions (C1).

5.6 Condition 6: Rational/Static View of the System Development Process

The sixth condition critiques the common perception of system development as a linear process, highlighting how this perspective overlooks the dynamic nature of organisations and the complex, often unpredictable ways individuals respond to technology over time (Bostrom & Heinen 1977a). It explains that socio-technical failures may not be evident at launch but can emerge and escalate after deployment.

This temporal and contextual perspective provides a new lens for understanding *bias*. A static view may overlook concept drift (i.e., data patterns change over time), leading to a decline in the performance of a once-accurate model (Bauer et al. 2024). More critically, it fails to recognise that a model's own outputs can influence its environment. Consequently, bias is not merely a static flaw but can evolve into a dynamic, self-reinforcing process. In this scenario, feedback loops can exacerbate initial unfairness as an unsupervised system generates increasingly skewed data.

This inability to foresee emergent human and systemic behaviours also sheds light on other gradually developing obstacles. For instance, *overreliance* often stems from "automation complacency" (Parasuraman & Manzey 2010), where users' vigilance diminishes due to prolonged, successful system use (Fleiss et al. 2024; Klingbeil et al. 2024). Such a static approach can also lead to mission creep, where systems for *algorithmic control* or data collection expand beyond their original intent, with increasing implications for *privacy* (Seoane & Velasco 2024). Similarly, a static view can solidify initial shifts in power into rigid structures over time, resulting in entrenched *organisational power dynamics* and imbalanced *global power dynamics*.

Ultimately, a static approach neglects the complex reality that trust and human behaviour are not fixed but constantly evolving. This creates a core dilemma between stability and change: Developers may fear the instability and costs associated with ongoing fairness interventions (von Zahn et al., 2022). Ironically, failing to accommodate unexpected changes can lead to even greater performance instability later when the system no longer serves its intended purpose (Akter, Dwivedi, et al, 2021; Buyl & de Bie 2024; Pessach & Shmueli 2021).

5.7 Condition 7: Limited Change Management

The final, seventh condition highlights a key pitfall that arises from the misconception that implementing a system guarantees success. This belief often results in minimal investment in essential areas like refining workflows, reorganising procedures, and addressing cultural factors crucial for effective adoption (Bostrom & Heinen 1977a). While the earlier conditions focus on flaws in the system's design, this one emphasises failures in the rollout and integration of technology within the social fabric of the organisation.

This perspective is vital when addressing user resistance and mistrust, such as *algorithmic aversion*. A limited change approach might assume that simply increasing technical transparency will build trust or fairness (Mirbabaie et al., 2025). However, literature shows that trust is also based on users' feeling respected and included throughout the process, which requires active communication and tailored support (Bankins et al. 2022; Siering 2022). Similarly, combating *overreliance* requires more than a single training session. A simple reminder of accountability does not significantly reduce users' inclination to follow flawed AI recommendations (Ghasemaghaei & Kordzadeh 2024b). This suggests that organisational processes need to be redesigned to genuinely promote accountability.

Effective change management is crucial for addressing structural barriers. For instance, bridging the *digital divide* is a specific change management task. When AI systems are implemented without sufficient support and resources for less tech-savvy individuals, it often leads to frustration and disengagement (Hu et al. 2024; S. Kim 2025). Similarly, the issue of *job displacement* can be addressed in various ways; its societal effects can be mitigated by actively providing training for new roles or redesigning jobs to work alongside AI, rather than simply replacing human labour (Marabelli et al. 2021).

Addressing concerns regarding breach of *privacy* also goes beyond mere technical compliance. It involves building trust through careful rollout processes that clearly explain how data is used and ensure that users have control over their data (Floridi et al. 2018; Liu & Tao 2022; Someh et al. 2019).

Condition 7 is central to the STS perspective since the analysis shows that many barriers can be traced to insufficient integration of the human element during implementation. Whether it's trust (users weren't onboarded properly), misuse (lack of training/policies), or resistance (stakeholders not included or considered), these are issues of change management. Fortunately, they are manageable if recognised. We can train, we can involve, we can set rules, but only if we plan for it. Plus, tensions become apparent: Organisations may rapidly deploy AI to reap benefits (speed as a goal, related to C4), but substantial change requires time and resources. Skipping the latter can backfire, for one, a botched implementation that needs revision or causes harm. Another tension is technology vs. user capacity: A modern AI might be state-of-the-art, but if users can't absorb it, its value diminishes. Therefore, effective change management should either make usage easier or elevate users to the needed capacity.

6 Conclusion and Summary

In this article, we argued that scholarship that deals with issues surrounding AI fails to adequately attend to the root causes. Consequently, solutions are insufficient given the ubiquitous implementation of AI. Since this pattern parallels Bostrom and Heinen's STS framework, we apply it to the contemporary AI corpus to shift the scholarly attention to the underlying causes of known issues. This shows how recurring challenges follow identifiable patterns linked to specific conditions that can be anticipated and, crucially, managed. To this end, we synthesised the literature into four thematic clusters: *Structural Inequality, Erosion of Trust, Erosion of Meaningful Work*, and *Consolidation of Power*. We then link their underlying dynamics to Bostrom and Heinen's seven conditions for MIS problems and failures.

Crucially, the seven conditions are interdependent. We observe recurrent couplings (e.g., a limited goal orientation (C4) narrows the analytic scope and reproduces a non-systemic view (C3)). These couplings surface design trade-offs that must be negotiated up front in AI governance (efficiency vs. equity, fairness vs. privacy, transparency vs. trust), turning tensions into design-time choices rather than after-the-fact surprises. By specifying how goal conflicts arise from condition couplings and locating them to design-time levers, we foreground the power dimensions that connect conditions.

By shifting the focus from surface-level symptoms to underlying causal mechanisms, we provide a diagnostic framework (see Table 1) that links the conditions that cause problems to socio-technical mechanisms and organisational consequences. For each condition, we specify common failure signatures and the mechanisms through which problems may intensify over time to facilitate diagnostics and to anticipate AI-related risks before they escalate. We invite future scholarship to translate our proposed mapping and solution principles (see Table 1) into actionable design guidance and solutions.

This work relocates several debates typically framed as ethical questions from the margins of deployment to the core of system design. We sincerely believe scholars and practitioners—and ultimately everyone—can benefit from attending to the root causes of emerging AI failures.

7 References

Akter, S., Dwivedi, Y. K., Biswas, K., Michael, K., Bandara, R. J., & Sajib, S. 2021. "Addressing Algorithmic Bias in AI-Driven Customer Management." *JOURNAL OF GLOBAL INFORMATION MANAGEMENT*, *29*(6). https://doi.org/10.4018/JGIM.20211101.oa3

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. 2021. "Algorithmic bias in data-driven innovation in the age of AI." *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *60*. https://doi.org/10.1016/j.ijinfomgt.2021.102387

Alkhatib, A. 2021. "To Live in Their Utopia: Why Algorithmic Systems Create Absurd Outcomes." *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–9. https://doi.org/10.1145/3411764.3445740

Arora, A., Barrett, M., Lee, E., Oborn, E., & Prince, K. 2023. “Risk and the future of AI: Algorithmic bias, data colonialism, and marginalization.” *Information and Organization*, *33*(3), 100478. https://doi.org/10.1016/j.infoandorg.2023.100478

Asatiani, A., Malo, P., Nagbol, P., Penttinen, E., Rinta-Kahila, T., & Salovaara, A. 2021. “Sociotechnical Envelopment of Artificial Intelligence: An Approach to Organizational Deployment of Inscrutable Artificial Intelligence Systems.” *JOURNAL OF THE ASSOCIATION FOR INFORMATION SYSTEMS*, *22*(2), 325–352. https://doi.org/10.17705/1jais.00664

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. 2022. “Ethical framework for Artificial Intelligence and Digital technologies.” *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *62*. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Avci, U. 2024. “Students’ GAI Acceptance: Role of Demographics, Creative Mindsets, Anxiety, Attitudes.” *JOURNAL OF COMPUTER INFORMATION SYSTEMS*. https://doi.org/10.1080/08874417.2024.2386545

Aydin, E., Rahman, M., & Ozeren, E. 2023. “Does Industry 5.0 Reproduce Gender (In)equalities at Organisations? Understanding the Interaction of Human Resources and Software Development Teams in Supplying Human Capitals.” *INFORMATION SYSTEMS FRONTIERS*. https://doi.org/10.1007/s10796-023-10450-1

Balakrishnan, J., Dwivedi, Y. K., Hughes, L., & Boy, F. 2024. “Enablers and Inhibitors of AI-Powered Voice Assistants: A Dual-Factor Approach by Integrating the Status Quo Bias and Technology Acceptance Model.” *INFORMATION SYSTEMS FRONTIERS*, *26*(3, SI), 921–942. https://doi.org/10.1007/s10796-021-10203-y

Bankins, S., Formosa, P., Griep, Y., & Richards, D. 2022. “AI Decision Making with Dignity? Contrasting Workers’ Justice Perceptions of Human and AI Decision Making in a Human Resource Management Context.” *INFORMATION SYSTEMS FRONTIERS*, *24*(3, SI), 857–875. https://doi.org/10.1007/s10796-021-10223-8

Baudier, P., & de Boissieu, E. 2025. “Are social media robot influencers credible? A cross-continental analysis in a fashion context.” *COMPUTERS IN HUMAN BEHAVIOR*, *162*. https://doi.org/10.1016/j.chb.2024.108447

Bauer, K., Heigl, R., Hinz, O., & Kosfeld, M. 2024. “Feedback Loops in Machine Learning: A Study on the Interplay of Continuous Updating and Human Discrimination.” *JOURNAL OF THE ASSOCIATION FOR INFORMATION SYSTEMS*, *25*(4). https://doi.org/10.17705/1jais.00853

Bedemariam, R., & Wessel, J. L. 2023. “The roles of outcome and race on applicant reactions to AI systems.” *COMPUTERS IN HUMAN BEHAVIOR*, *148*. https://doi.org/10.1016/j.chb.2023.107869

Bednar, P., & Welch, C. 2020. “Socio-Technical Perspectives on Smart Working: Creating Meaningful and Sustainable Systems.” *INFORMATION SYSTEMS FRONTIERS*, *22*(2), 281–298. https://doi.org/10.1007/s10796-019-09921-1

Bigman, Y. E., Yam, K. C., Marciano, D., Reynolds, S. J., & Gray, K. 2021. “Threat of racial and economic inequality increases preference for algorithm decision-making.” *COMPUTERS IN HUMAN BEHAVIOR*, *122*. https://doi.org/10.1016/j.chb.2021.106859

Binesh, F., & Baloglu, S. 2023. “Are we ready for hotel robots after the pandemic? A profile analysis.” *COMPUTERS IN HUMAN BEHAVIOR*, *147*. https://doi.org/10.1016/j.chb.2023.107854

Birzhandi, P., & Cho, Y.-S. 2023. “Application of fairness to healthcare, organizational justice, and finance: A survey.” *EXPERT SYSTEMS WITH APPLICATIONS*, *216*. https://doi.org/10.1016/j.eswa.2022.119465

Boell, S., & Wang, B. 2019. “Www.litbaskets.io, an IT Artifact Supporting Exploratory Literature Searches for Information Systems Research.” *Proceedings of the 2019 Australasian Conference on Information Systems*. https://aisel.aisnet.org/acis2019/71

Bostrom, R. P., & Heinen, J. S. 1977a. „MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes.” *MIS QUARTERLY*, *1*(3), 17–32. https://doi.org/10.2307/248710

Bostrom, R. P., & Heinen, J. S. 1977b. “MIS Problems and Failures: A Socio-Technical Perspective, Part II: The Application of Socio-Technical Theory.” *MIS QUARTERLY*, *1*(4), 11–28. https://doi.org/10.2307/249019

Buyl, M., & de Bie, T. 2024. “Inherent Limitations of AI Fairness.” *COMMUNICATIONS OF THE ACM*, *67*(2), 48–55. https://doi.org/10.1145/3624700

Carter, L., Liu, D., & Cantrell, C. 2020. “Exploring the Intersection of the Digital Divide and Artificial Intelligence: A Hermeneutic Literature Review.” *AIS Transactions on Human-Computer Interaction*, *12*(4), 253–275. https://doi.org/10.17705/1thci.00138

Chang, R. C.-S., Lu, H.-P., & Yang, P. 2018. „Stereotypes or golden rules? Exploring likable voice traits of social robots as active aging companions for tech-savvy baby boomers in Taiwan.”

COMPUTERS IN HUMAN BEHAVIOR, *84*, 194–210. https://doi.org/10.1016/j.chb.2018.02.025

Chiu, Y., Zhu, Y., & Corbett, J. 2021. "In the hearts and minds of employees: A model of pre-adoptive appraisal toward artificial intelligence in organizations." *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *60*. https://doi.org/10.1016/j.ijinfomgt.2021.102379

Claure, H., Kim, S., Kizilcec, R. F., & Jung, M. 2023. "The social consequences of Machine Allocation Behavior: Fairness, interpersonal perceptions and performance." *COMPUTERS IN HUMAN BEHAVIOR*, *146*. https://doi.org/10.1016/j.chb.2022.107628

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. 2021. "Artificial intelligence in information systems research: A systematic literature review and research agenda." *International Journal of Information Management*, *60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Cornacchia, G., Anelli, V. W., Biancofiore, G. M., Narducci, F., Pomo, C., Ragone, A., & Di Sciascio, E. 2023. "Auditing fairness under unawareness through counterfactual reasoning." *INFORMATION PROCESSING & MANAGEMENT*, *60*(2). https://doi.org/10.1016/j.ipm.2022.103224

Cowgill, B., Dell'Acqua, F., & Matz, S. 2020. "The Managerial Effects of Algorithmic Fairness Activism." *AEA Papers and Proceedings*, *110*, 85–90. https://doi.org/10.1257/pandp.20201035

Di Noia, T., Tintarev, N., Fatourou, P., & Schedl, M. 2022. Recommender Systems under European AI Regulations. *COMMUNICATIONS OF THE ACM*, *65*(4), 69–73. https://doi.org/10.1145/3512728

Dimaggio, P., & Powell, W. W. 1983. "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields." *American Sociological Review*, *48*(2).

Duda, S., Hofmann, P., Urbach, N., Voelter, F., & Zwickel, A. 2024. "The Impact of Resource Allocation on the Machine Learning Lifecycle." *BUSINESS & INFORMATION SYSTEMS ENGINEERING*, *66*(2), 203–219. https://doi.org/10.1007/s12599-023-00842-7

Engstrom, A., Pittino, D., Mohlin, A., Johansson, A., & Mirzaei, N. E. 2024. "Artificial intelligence and work transformations: Integrating sensemaking and workplace learning perspectives." *INFORMATION TECHNOLOGY & PEOPLE*, *37*(7), 2441–2461. https://doi.org/10.1108/ITP-01-2023-0048

Feuerriegel, S., Dolata, M., & Schwabe, G. 2020. "Fair AI Challenges and Opportunities." *BUSINESS & INFORMATION SYSTEMS ENGINEERING*, *62*(4), 379–384. https://doi.org/10.1007/s12599-020-00650-3

Fischer, L. H., & Baskerville, R. 2023. "Explaining sociotechnical change: An unstable equilibrium perspective." *EUROPEAN JOURNAL OF INFORMATION SYSTEMS*, *32*(4), 634–652. https://doi.org/10.1080/0960085X.2021.2023669

Fleiss, J., Back, E., & Thalmann, S. 2024. "Mitigating Algorithm Aversion in Recruiting: A Study on Explainable AI for Conversational Agents." *DATA BASE FOR ADVANCES IN INFORMATION SYSTEMS*, *55*(1), 56–87. https://doi.org/10.1145/3645057.3645062

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. 2018. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *MINDS AND MACHINES*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Fosso Wamba, S., Queiroz, M. M., Pappas, I. O., & Sullivan, Y. 2024. "Artificial Intelligence Capability and Firm Performance: A Sustainable Development Perspective by the Mediating Role of Data-Driven Culture." *INFORMATION SYSTEMS FRONTIERS*. https://doi.org/10.1007/s10796-023-10460-z

Ghasemaghaei, M., & Kordzadeh, N. 2024a. "Ethics in the Age of Algorithms: Unravelling the Impact of Algorithmic Unfairness on Data Analytics Recommendation Acceptance." *INFORMATION SYSTEMS JOURNAL*. https://doi.org/10.1111/isj.12572

Ghasemaghaei, M., & Kordzadeh, N. 2024b. "Understanding how algorithmic injustice leads to making discriminatory decisions: An obedience to authority perspective." *INFORMATION & MANAGEMENT*, *61*(2). https://doi.org/10.1016/j.im.2024.103921

Gravett, W. 2022. "Digital Neocolonialism: The Chinese Surveillance State in Africa." *AFRICAN JOURNAL OF INTERNATIONAL AND COMPARATIVE LAW*, *30*, 39–58. https://doi.org/10.3366/ajicl.2022.0393

Gude, V. 2023. "Factors Influencing ChatGpt Adoption for Product Research and Information Retrieval." *JOURNAL OF COMPUTER INFORMATION SYSTEMS*. https://doi.org/10.1080/08874417.2023.2280918

Gupta, M., Parra, C. M., & Dennehy, D. 2022. "Questioning Racial and Gender Bias in AI-based Recommendations: Do Espoused National Cultural Values Matter?" *INFORMATION SYSTEMS FRONTIERS*, *24*(5), 1465–1481. https://doi.org/10.1007/s10796-021-10156-2

Gyldenkaerne, C., Hansen, J. U., Hertzum, M., & Monsted, T. 2024. “Innovation tactics for implementing an ML application in healthcare: A long and winding road.” *INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES*, *181*. https://doi.org/10.1016/j.ijhcs.2023.103162

Heeks, R., Ospina, A. V., Foster, C., Gao, P., Han, X., Jepson, N., Schindler, S., & Zhou, Q. 2024. “China’s digital expansion in the Global South: Systematic literature review and future research agenda.” *INFORMATION SOCIETY*, *40*(2, SI), 69–95. https://doi.org/10.1080/01972243.2024.2315875

Heng, S., Tsilionis, K., Scharff, C., & Wautelet, Y. 2022. “Understanding AI ecosystems in the Global South: The cases of Senegal and Cambodia.” *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *64*. https://doi.org/10.1016/j.ijinfomgt.2021.102454

Hoddinghaus, M., Sondern, D., & Hertel, G. 2021. “The automation of leadership functions: Would people trust decision algorithms?” *COMPUTERS IN HUMAN BEHAVIOR*, *116*. https://doi.org/10.1016/j.chb.2020.106635

Hoffmann, A. L. 2019. “Where fairness fails: Data, algorithms, and the limits of antidiscrimination discourse.” *INFORMATION, COMMUNICATION & SOCIETY*, *22*(7), 900–915. https://doi.org/10.1080/1369118X.2019.1573912

Hong, J.-W., Fischer, K., Ha, Y., & Zeng, Y. 2022. “Human, I wrote a song for you: An experiment testing the influence of machines’ attributes on the AI-composed music evaluation.” *COMPUTERS IN HUMAN BEHAVIOR*, *131*. https://doi.org/10.1016/j.chb.2022.107239

Hoorn, J. F., & Huang, I. S. 2024. “The media inequality, uncanny mountain, and the singularity is far from near: Iwaa and Sophia robot versus a real human being.” *INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES*, *181*. https://doi.org/10.1016/j.ijhcs.2023.103142

Hu, P., Zeng, Y., Wang, D., & Teng, H. 2024. “Too much light blinds: The transparency-resistance paradox in algorithmic management.” *COMPUTERS IN HUMAN BEHAVIOR*, *161*. https://doi.org/10.1016/j.chb.2024.108403

Isaac, M. S., Wang, R. J.-H., Napper, L. E., & Marsh, J. K. 2024. “To err is human: Bias salience can help overcome resistance to medical AI.” *COMPUTERS IN HUMAN BEHAVIOR*, *161*. https://doi.org/10.1016/j.chb.2024.108402

Jahangir, R., Teh, Y. W., Nweke, H. F., Mujtaba, G., Al-Garadi, M. A., & Ali, I. 2021. “Speaker identification through artificial intelligence techniques: A comprehensive review and research challenges.” *EXPERT SYSTEMS WITH APPLICATIONS*, *171*. https://doi.org/10.1016/j.eswa.2021.114591

Jeon, S., Chang, Y. S., Jo, S. J., Madukuand, T., & Kim, Y. E. 2024. “Speed of Catch-Up and Convergence of the Artificial Intelligence Divide: AI Investment, Robotic, Start-Ups, and Patents.” *JOURNAL OF GLOBAL INFORMATION TECHNOLOGY MANAGEMENT*, *27*(1), 63–85. https://doi.org/10.1080/1097198X.2023.2297636

Keyes, O. 2018. The Misgendering Machines: “Trans/HCI Implications of Automatic Gender Recognition.” *Proceedings of the ACM on Human-Computer Interaction*. https://doi.org/10.1145/3274357

Kim, D., & Song, H. 2024. “Designing an age-friendly conversational AI agent for mobile banking: The effects of voice modality and lip movement.” *INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES*, *187*. https://doi.org/10.1016/j.ijhcs.2024.103262

Kim, S. 2025. “Perceptions of discriminatory decisions of artificial intelligence: Unpacking the role of individual characteristics.” *INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES*, *194*. https://doi.org/10.1016/j.ijhcs.2024.103387

Klingbeil, A., Gruetzner, C., & Schreck, P. 2024. “Trust and reliance on AI - An experimental study on the extent and costs of overreliance on AI.” *COMPUTERS IN HUMAN BEHAVIOR*, *160*. https://doi.org/10.1016/j.chb.2024.108352

Koechling, A., Riazy, S., Wehner, M. C., & Simbeck, K. 2021. “Highly Accurate, But Still Discriminatory A Fairness Evaluation of Algorithmic Video Analysis in the Recruitment Context.” *BUSINESS & INFORMATION SYSTEMS ENGINEERING*, *63*(1), 39–54. https://doi.org/10.1007/s12599-020-00673-w

Kordzadeh, N., & Ghasemaghaei, M. 2022. “Algorithmic bias: Review, synthesis, and future research directions.” *EUROPEAN JOURNAL OF INFORMATION SYSTEMS*, *31*(3), 388–409. https://doi.org/10.1080/0960085X.2021.1927212

Kushwaha, A. K., Pharswan, R., Kumar, P., & Kar, A. K. 2023. “How Do Users Feel When They Use Artificial Intelligence for Decision Making? A Framework for Assessing Users’ Perception.” *INFORMATION SYSTEMS FRONTIERS*, *25*(3, SI), 1241–1260. https://doi.org/10.1007/s10796-022-10293-2

Kwet, M. 2019. "Digital colonialism: US empire and the new imperialism in the Global South." *RACE & CLASS*, *60*(4), 3–26. https://doi.org/10.1177/0306396818823172

Laine, J., Minkkinen, M., & Mantymaki, M. 2024. "Ethics-based AI auditing: A systematic literature review on conceptualizations of ethical principles and knowledge contributions to stakeholders." *INFORMATION & MANAGEMENT*, *61*(5). https://doi.org/10.1016/j.im.2024.103969

Lambrecht, A., & Tucker, C. 2019. "Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads." *MANAGEMENT SCIENCE*, *65*(7), 2966–2981. https://doi.org/10.1287/mnsc.2018.3093

Langer, M., & Landers, R. N. 2021. "The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers." *COMPUTERS IN HUMAN BEHAVIOR*, *123*. https://doi.org/10.1016/j.chb.2021.106878

Liu, K., & Tao, D. 2022. "The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services." *COMPUTERS IN HUMAN BEHAVIOR*, *127*. https://doi.org/10.1016/j.chb.2021.107026

Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. 2021. "Artificial intelligence in business: State of the art and future research agenda." *JOURNAL OF BUSINESS RESEARCH*, *129*, 911–926. https://doi.org/10.1016/j.jbusres.2020.11.001

Lu, T., & Zhang, Y. 2024. "1+1 >2? Information, Humans, and Machines." *INFORMATION SYSTEMS RESEARCH*. https://doi.org/10.1287/isre.2023.0305

Lu, Y. 2019. "Artificial intelligence: A survey on evolution, models, applications and future trends." *JOURNAL OF MANAGEMENT ANALYTICS*, *6*(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. 2020. "Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization." *JOURNAL OF BUSINESS RESEARCH*, *120*, 262–273. https://doi.org/10.1016/j.jbusres.2020.07.045

Marabelli, M., Vaast, E., & Li, J. L. 2021. "Preventing the digital scars of COVID-19." *EUROPEAN JOURNAL OF INFORMATION SYSTEMS*, *30*(2), 176–192. https://doi.org/10.1080/0960085X.2020.1863752

Marjanovic, O., Cecez-Kecmanovic, D., & Vidgen, R. 2022. "Theorising Algorithmic Justice." *EUROPEAN JOURNAL OF INFORMATION SYSTEMS*, *31*(3), 269–287. https://doi.org/10.1080/0960085X.2021.1934130

Marotta, A. 2022. "When AI Is Wrong: Addressing Liability Challenges in Women's Healthcare." *JOURNAL OF COMPUTER INFORMATION SYSTEMS*, *62*(6), 1310–1319. https://doi.org/10.1080/08874417.2022.2089773

Martin, K. 2019. "Ethical Implications and Accountability of Algorithms." *JOURNAL OF BUSINESS ETHICS*, *160*(4), 835–850. https://doi.org/10.1007/s10551-018-3921-3

Masiero, S., & Das, S. 2019. "Datafying anti-poverty programmes: Implications for data justice." *INFORMATION, COMMUNICATION & SOCIETY*, *22*(7), 916–933. https://doi.org/10.1080/1369118X.2019.1575448

Merhi, M. 2023. "An Assessment of the Barriers Impacting Responsible Artificial Intelligence." *INFORMATION SYSTEMS FRONTIERS*, *25*(3), 1147–1160. https://doi.org/10.1007/s10796-022-10276-3

Mirbabaie, M., Brünker, F., Möllmann Frick, N. R. J., & Stieglitz, S. 2022. "The rise of artificial intelligence – understanding the AI identity threat at the workplace." *ELECTRONIC MARKETS*, *32*(1), 73–99. https://doi.org/10.1007/s12525-021-00496-x

Mirbabaie, M., Langer, M., Rieskamp, J., and Hofeditz, L. 2025. "Transparency Fallacy: Perceived Fairness in Algorithmic Management," BUSINESS & INFORMATION SYSTEMS ENGINEERING. https://doi.org/10.1007/s12599-025-00963-1.

Mou, Y., & Xu, K. 2017. "The media inequality: Comparing the initial human-human and human-AI social interactions." *COMPUTERS IN HUMAN BEHAVIOR*, *72*, 432–440. https://doi.org/10.1016/j.chb.2017.02.067

Moussawi, S., Deng, X. (Nancy), & Joshi, K. D. 2024. "AI and Discrimination: Sources of Algorithmic Biases." *DATA BASE FOR ADVANCES IN INFORMATION SYSTEMS*, *55*(4), 6–11. https://doi.org/10.1145/3701613.3701615

Nedungadi, P., Ramesh, M., Govindaraju, V., Rao, B., Berbeglia, P., & Raman, R. 2024. "Emerging leaders or persistent gaps? Generative AI research may foster women in STEM." *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *77*. https://doi.org/10.1016/j.ijinfomgt.2024.102785

Noor, N., Hill, S. R., & Troshani, I. 2022. “Artificial Intelligence Service Agents: Role of Parasocial Relationship.” *JOURNAL OF COMPUTER INFORMATION SYSTEMS*, *62*(5), 1009–1023. https://doi.org/10.1080/08874417.2021.1962213

Ochmann, J., Michels, L., Tiefenbeck, V., Maier, C., & Laumer, S. 2024. “Perceived algorithmic fairness: An empirical study of transparency and anthropomorphism in algorithmic recruiting.” *INFORMATION SYSTEMS JOURNAL*, *34*(2), 384–414. https://doi.org/10.1111/isj.12482

Oleksy, T., Wnuk, A., Domaradzka, A., & Maison, D. 2023. “What shapes our attitudes towards algorithms in urban governance? The role of perceived friendliness and controllability of the city, and human-algorithm cooperation.” *COMPUTERS IN HUMAN BEHAVIOR*, *142*. https://doi.org/10.1016/j.chb.2023.107653

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., and Moher, D. 2021. “The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews,” *SYSTEMATIC REVIEWS, (10)1*. https://doi.org/10.1186/s13643-021-01626-4.

Parasuraman, R., & Manzey, D. H. 2010. “Complacency and Bias in Human Use of Automation: An Attentional Integration.” *HUMAN FACTORS*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Pessach, D., & Shmueli, E. 2021. “Improving fairness of artificial intelligence algorithms in Privileged-Group Selection Bias data settings.” *EXPERT SYSTEMS WITH APPLICATIONS*, *185*. https://doi.org/10.1016/j.eswa.2021.115667

Porra, J., Lacity, M., & Parks, M. S. 2020. “Can Computer Based Human-Likeness Endanger Humanness?”—A Philosophical and Ethical Perspective on Digital Assistants Expressing Feelings They Can’t Have.” *INFORMATION SYSTEMS FRONTIERS*, *22*(3, SI), 533–547. https://doi.org/10.1007/s10796-019-09969-z

Qin, S., Jia, N., Luo, X., Liao, C., & Huang, Z. 2023. “Perceived Fairness of Human Managers Compared with Artificial Intelligence in Employee Performance Evaluation.” *JOURNAL OF MANAGEMENT INFORMATION SYSTEMS*, *40*(4), 1039–1070. https://doi.org/10.1080/07421222.2023.2267316

Raisch, S., & Fomina, K. 2025. “Combining Human and Artificial Intelligence: Hybrid Problem-Solving in Organizations.” *ACADEMY OF MANAGEMENT REVIEW*, *50*(2), 441–464. https://doi.org/10.5465/amr.2021.0421

Ransbotham, S., Fichman, R. G., Gopal, R., & Gupta, A. (2016). “Special Section Introduction-Ubiquitous IT and Digital Vulnerabilities.” *INFORMATION SYSTEMS RESEARCH*, *27*(4), 834–847. https://doi.org/10.1287/isre.2016.0683

Rhue, L. 2024. “The Anchoring Effect, Algorithmic Fairness, and the Limits of Information Transparency for Emotion Artificial Intelligence.” *INFORMATION SYSTEMS RESEARCH*, *35*(3), 1479–1496. https://doi.org/10.1287/isre.2019.0493

Rieskamp, J., Mirbabaie, M., and Zander, K. 2023. “GenAI-Powered Social Bots for Crisis Communication: A Systematic Literature Review,” in PROCEEDINGS OF THE 2023 AUSTRALASIAN CONFERENCE ON INFORMATION SYSTEMS, Wellington. https://aisel.aisnet.org/acis2023/65.

Rondina, M., Vinci, F., Vetrò, A., & De Martin, J. C. 2025. “Facial Analysis Systems and Down Syndrome.” In R. Meo & F. Silvestri (Eds.), *Machine Learning and Principles and Practice of Knowledge Discovery in Databases* (pp. 145–160). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-74630-7_10

Russo, P. A., Duradoni, M., & Guazzini, A. 2021. “How self-perceived reputation affects fairness towards humans and artificial intelligence.” *COMPUTERS IN HUMAN BEHAVIOR*, *124*. https://doi.org/10.1016/j.chb.2021.106920

Sahu, S., & Mohanty, A. 2024. “What’s Next for U.S. AI Policy? | Carnegie Endowment for International Peace.” (https://carnegieendowment.org/posts/2024/12/whats-next-for-us-ai-policy?lang=en, accessed August 28th, 2025)

Sartori, L., & Theodorou, A. 2022. “A sociotechnical perspective for the future of AI: Narratives, inequalities, and human control.” *ETHICS AND INFORMATION TECHNOLOGY*, *24*(1), 4. https://doi.org/10.1007/s10676-022-09624-3

Schanke, S., Burtch, G., & Ray, G. 2021. “Estimating the Impact of “Humanizing” Customer Service Chatbots.” *INFORMATION SYSTEMS RESEARCH*, *32*(3), 736–751. https://doi.org/10.1287/isre.2021.1015

Schneider, J., Abraham, R., Meske, C., & Vom Brocke, J. 2023. “Artificial Intelligence Governance For Businesses.” *INFORMATION SYSTEMS MANAGEMENT*, *40*(3), 229–249. https://doi.org/10.1080/10580530.2022.2085825

Seoane, M. F. V., & Velasco, C. M. A. 2024. “The Chinese surveillance state in Latin America? Evidence from Argentina and Ecuador.” *INFORMATION SOCIETY*, *40*(2, SI), 154–167. https://doi.org/10.1080/01972243.2024.2317057

Sestino, A., & De Mauro, A. 2022. “Leveraging Artificial Intelligence in Business: Implications, Applications and Methods.” *TECHNOLOGY ANALYSIS & STRATEGIC MANAGEMENT*, *34*(1), 16–29. https://doi.org/10.1080/09537325.2021.1883583

Sharma, M., Joshi, S., Luthra, S., & Kumar, A. 2024. “Impact of Digital Assistant Attributes on Millennials’ Purchasing Intentions: A Multi-Group Analysis using PLS-SEM, Artificial Neural Network and fsQCA.” *INFORMATION SYSTEMS FRONTIERS*, *26*(3, SI), 943–966. https://doi.org/10.1007/s10796-022-10339-5

Sharma, P., Gunasekaran, A., & Subramanian, G. 2024. “Enhancing Supply Chain: Exploring and Exploiting AI Capabilities.” *JOURNAL OF COMPUTER INFORMATION SYSTEMS*. https://doi.org/10.1080/08874417.2024.2386527

Siering, M. 2022. “Explainability and fairness of RegTech for regulatory enforcement: Automated monitoring of consumer complaints.” *DECISION SUPPORT SYSTEMS*, *158*. https://doi.org/10.1016/j.dss.2022.113782

Sipior, J., Ward, B., Rusinko, C., & Lombardi, D. 2024. “Bias in Using AI for Recruiting: Legal Considerations.” *INFORMATION SYSTEMS MANAGEMENT*, *41*(4), 399–412. https://doi.org/10.1080/10580530.2023.2294453

Someh, I., Davern, M., Breidbach, C. F., & Shanks, G. 2019. “Ethical Issues in Big Data Analytics: A Stakeholder Perspective.” *COMMUNICATIONS OF THE ASSOCIATION FOR INFORMATION SYSTEMS*, 44, 718–747. https://doi.org/10.17705/1CAIS.04434

Thomas, A. 2024. “Digitally transforming the organization through knowledge management: A socio-technical system (STS) perspective.” *EUROPEAN JOURNAL OF INNOVATION MANAGEMENT*, *27*(9), 437–460. https://doi.org/10.1108/EJIM-02-2024-0114

Trist, E. L., & Bamforth, K. W. (1951). “Some Social and Psychological Consequences of the Longwall Method of Coal-Getting: An Examination of the Psychological Situation and Defences of a Work Group in Relation to the Social Structure and Technological Content of the Work System.” *HUMAN RELATIONS*, *4*(1), 3–38. https://doi.org/10.1177/001872675100400101

Tsung-Yu, H., Yu-Chia, T., & Wen (Tina), Y. C. 2024. “Is this AI sexist? The effects of a biased AI’s anthropomorphic appearance and explainability on users’ bias perceptions and trust.” *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *76*. https://doi.org/10.1016/j.ijinfomgt.2024.102775

van Berkel, N., Sarsenbayeva, Z., & Goncalves, J. 2023. “The methodology of studying fairness perceptions in Artificial Intelligence: Contrasting CHI and FAccT.” *INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES*, *170*. https://doi.org/10.1016/j.ijhcs.2022.102954

von Zahn, M., Feuerriegel, S., & Kuehl, N. 2022. “The Cost of Fairness in AI: Evidence from E-Commerce.” *BUSINESS & INFORMATION SYSTEMS ENGINEERING*, *64*(3), 335–348. https://doi.org/10.1007/s12599-021-00716-w

Wang, C., Teo, T. S. H., & Janssen, M. 2021. “Public and private value creation using artificial intelligence: An empirical study of AI voice robot users in Chinese public sector.” *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *61*. https://doi.org/10.1016/j.ijinfomgt.2021.102401

Weber, M., Engert, M., Schaffer, N., Weking, J., & Krcmar, H. 2023. “Organizational Capabilities for AI Implementation-Coping with Inscrutability and Data Dependency in AI.” *INFORMATION SYSTEMS FRONTIERS*, *25*(4), 1549–1569. https://doi.org/10.1007/s10796-022-10297-y

Wesche, J. S., & Sonderegger, A. 2021. “Repelled at first sight? Expectations and intentions of job-seekers reading about AI selection in job advertisements.” *COMPUTERS IN HUMAN BEHAVIOR*, *125*. https://doi.org/10.1016/j.chb.2021.106931

Wu, M., Li, Z., & Yuen, K. F. 2024. “Effect of anthropomorphic design and hierarchical status on balancing self-serving bias: Accounting for education, ethnicity, and experience.” *COMPUTERS IN HUMAN BEHAVIOR*, *158*. https://doi.org/10.1016/j.chb.2024.108299

Yan, J., Leidner, D. E., & Peters, U. 2024. “Global techno-politics: A review of the current status and opportunities for future research.” *INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT*, *75*. https://doi.org/10.1016/j.ijinfomgt.2023.102729

Yanit, M., Yanit, M., & Wan, F. 2023. “Right agent, wrong level of hedonism: How high (vs low) hedonic values in AI-performed tasks lead to decreased perceptions of humanlikeness, warmth, and less consumer support.” *COMPUTERS IN HUMAN BEHAVIOR*, *147*. https://doi.org/10.1016/j.chb.2023.107870

Yarger, L., Cobb Payton, F., & Neupane, B. 2020. “Algorithmic equity in the hiring of underrepresented IT job candidates.” *ONLINE INFORMATION REVIEW*, *44*(2), 383–395. https://doi.org/10.1108/OIR-10-2018-0334

Zhang, D., Peng, G., Yao, Y., & Browning, T. R. 2024. “Is a College Education Still Enough? The IT-Labor Relationship with Education Level, Task Routineness, and Artificial Intelligence.” *INFORMATION SYSTEMS RESEARCH*, *35*(3), 992–1010. https://doi.org/10.1287/isre.2021.0391

Zhdanov, D., Bhattacharjee, S., & Bragin, M. A. 2022. “Incorporating FAT and privacy aware AI modeling approaches into business decision making frameworks.” *DECISION SUPPORT SYSTEMS*, *155*. https://doi.org/10.1016/j.dss.2021.113715

Appendix 1: Conditions of the Socio-Technical Systems Perspective

Condition	Core Idea
C1: Systems Designer's Implicit Theories	Designers' unexamined, implicit assumptions about human behaviour (e.g., viewing users as inefficient or change-resistant) unconsciously shape the system design, often leading to alienating outcomes.
C2: Systems Designer's Concept of Responsibility	A flawed concept of responsibility where designers assume full control over the change, failing to empower users to retain ownership, which is a prerequisite for meaningful adoption and success.
C3: Non-Systemic View	A narrow focus on the immediate technical system while ignoring the secondary effects or ripple effects on the broader social structure, work relationships, and distribution of authority.
C4: Limited Goal Orientation	A narrow definition of success focused solely on technical or economic goals (e.g., efficiency), while crucial social objectives like quality of working life (QWL) are seen as secondary or sacrificed.
C5: Limited Design Referent Group	The system is designed for a narrow group of primary users (e.g., management), while the needs and roles of secondary or peripheral users (e.g., clerks, data entry staff) are ignored, despite their critical role in the system's overall performance.
C6: Rational/Static View of Development	The idealized view of system development as a linear, rational sequence of steps, which fails to account for the dynamic, political, and iterative nature of organizational change and stakeholder negotiations.
C7: Limited Change Technologies	An exclusive reliance on technical interventions and basic training, while neglecting the broader toolkit from behavioural sciences and Organizational Development (OD) needed to address social system issues like job redesign or interdepartmental conflict.

Table 2 The Seven Conditions Leading to MIS Failures (based on Bostrom & Heinen, 1977a)

Appendix 2: Literature Search and Mapping

Database	Search Query
Web of Science	TS=("Artificial Intelligence" OR "AI" OR "LLM" OR "Large Language Model*") AND TS=("fair*" OR "transparen*" OR "algorithmic bias" OR "sexis*" OR "gender" OR "racis*" OR "digital divide" OR "inequal*" OR "global south" OR "*colonial*")
AISel	abstract:("fair*" OR "transparen*" OR "algorithmic bias" OR "sexis*" OR "gender" OR "racis*" OR "digital divide" OR "inequal*" OR "global south" OR "*colonial*") AND abstract:("artificial intelligence" OR "AI" OR "LLM" OR "large language model*“)

Table 3 Structured Search Query for Database Retrieval

Concept	**Referenced by**
Reinforcement of Structural Inequality	
Bias	(Akter, Dwivedi, et al., 2021; Akter, McCarthy, et al., 2021; Alkhatib, 2021; Arora et al., 2023; Aydin et al., 2023; Bauer et al., 2024; Bedemariam & Wessel, 2023; Bigman et al., 2021; Birzhandi & Cho, 2023; Buyl & de Bie, 2024; Carter et al., 2020; Cornacchia et al., 2023; Cowgill et al., 2020; Feuerriegel et al., 2020; Fleiss et al., 2024; Floridi et al., 2018; Ghasemaghaei & Kordzadeh, 2024a, 2024b; Gupta et al., 2022; Hoffmann, 2019; Isaac et al., 2024; Jahangir et al., 2021; Keyes, 2018; S. Kim, 2025; Koechling et al., 2021; Kordzadeh & Ghasemaghaei, 2022; Kushwaha et al., 2023; Laine et al., 2024; Lambrecht & Tucker, 2019; Langer & Landers, 2021; Loureiro et al., 2021, 2021; Y. Lu, 2019; Marabelli et al., 2021; Marjanovic et al., 2022; Marotta, 2022; Martin, 2019; Ochmann et al., 2024; Pessach & Shmueli, 2021; Ransbotham et al., 2016; Rhue, 2024; Rondina et al., 2025; Siering, 2022; Sipior et al., 2024; Tsung-Yu et al., 2024; van Berkel et al., 2023, 2023; von Zahn et al., 2022; Wesche & Sonderegger, 2021; Yan et al., 2024; Yarger et al., 2020; Zhdanov et al., 2022)
Age	(Birzhandi & Cho, 2023; Gupta et al., 2022; Isaac et al., 2024; S. Kim, 2025; Siering, 2022)
Gender	(Alkhatib, 2021; Aydin et al., 2023; Bauer et al., 2024; Birzhandi & Cho, 2023; Buyl & de Bie, 2024; Cornacchia et al., 2023; Feuerriegel et al., 2020; Ghasemaghaei & Kordzadeh, 2024a; Gupta et al., 2022; Hoffmann, 2019; Isaac et al., 2024; Keyes, 2018; Koechling et al., 2021; Kordzadeh & Ghasemaghaei, 2022; Kushwaha et al., 2023; Lambrecht & Tucker, 2019; Marjanovic et al., 2022; Marotta, 2022; Moussawi et al., 2024; Nedungadi et al., 2024; Pessach & Shmueli, 2021; Russo et al., 2021; Tsung-Yu et al., 2024; Wang et al., 2021; Yarger et al., 2020)
Transgender	(Alkhatib, 2021; Keyes, 2018)
Ethnicity	(Alkhatib, 2021; Bedemariam & Wessel, 2023; Bigman et al., 2021; Birzhandi & Cho, 2023; Buyl & de Bie, 2024; Cornacchia et al., 2023; Feuerriegel et al., 2020; Gupta et al., 2022; Hoffmann, 2019; Isaac et al., 2024; Koechling et al., 2021; Kordzadeh & Ghasemaghaei, 2022; Kushwaha et al., 2023; Marabelli et al., 2021; Marjanovic et al., 2022; Martin, 2019; Moussawi et al., 2024; Pessach & Shmueli, 2021; Rhue, 2024; Yarger et al., 2020)
Disability	(Birzhandi & Cho, 2023; Buyl & de Bie, 2024; Hoffmann, 2019; Jahangir et al., 2021; Marjanovic et al., 2022; Rondina et al., 2025)
Digital Divide	(Arora et al., 2023; Binesh & Baloglu, 2023; Carter et al., 2020; Heng et al., 2022; Hu et al., 2024; Jeon et al., 2024; D. Kim & Song, 2024; S. Kim, 2025; Loureiro et al., 2021; Marabelli et al., 2021; Marjanovic et al., 2022; Ransbotham et al., 2016; P. Sharma et al., 2024; Wang et al., 2021; Yan et al., 2024)
Erosion of Trust	
Anthropomorphism/Uncanny Valley	(Baudier & de Boissieu, 2025; Binesh & Baloglu, 2023; Chang et al., 2018; Hong et al., 2022; Hoorn & Huang, 2024; Hu et al., 2024; Liu & Tao, 2022; Mou & Xu, 2017; Noor et al., 2022; Ochmann et al., 2024; Oleksy et al., 2023; Porra et al., 2020; Russo et al., 2021; Schanke et al., 2021; M. Sharma et al., 2024; Sipior et al., 2024; Tsung-Yu et al., 2024; Wu et al., 2024; Yanit et al., 2023)
Dehumanisation	(Bankins et al., 2022; Bedemariam & Wessel, 2023; Bigman et al., 2021; Floridi et al., 2018; Kwet, 2019; Langer & Landers, 2021; Loureiro et al., 2021; Y. Lu, 2019; Marjanovic et al., 2022; Marotta, 2022; Porra et al., 2020; van Berkel et al., 2023; Wesche & Sonderegger, 2021)
Breach of Privacy	(Balakrishnan et al., 2024; Floridi et al., 2018; Gravett, 2022; Kushwaha et al., 2023; Kwet, 2019; Laine et al., 2024; Langer & Landers, 2021; Liu & Tao, 2022; Y. Lu, 2019; Marabelli et al., 2021; Marjanovic et al., 2022; Marotta, 2022; Masiero & Das, 2019; Ransbotham et al., 2016; Seoane & Velasco, 2024; Sipior et al., 2024; Someh et al., 2019; Zhdanov et al., 2022)

Algorithmic Aversion	(Avci, 2024; Balakrishnan et al., 2024; Bankins et al., 2022; Baudier & de Boissieu, 2025; Bauer et al., 2024; Bedemariam & Wessel, 2023; Bigman et al., 2021; Binesh & Baloglu, 2023; Birzhandi & Cho, 2023; Chang et al., 2018; Feuerriegel et al., 2020; Fleiss et al., 2024; Floridi et al., 2018; Ghasemaghaei & Kordzadeh, 2024a, 2024b; Gude, 2023; Gupta et al., 2022; Hoddinghaus et al., 2021; Hong et al., 2022; Hoorn & Huang, 2024; Hu et al., 2024; Isaac et al., 2024; D. Kim & Song, 2024; S. Kim, 2025; Klingbeil et al., 2024; Koechling et al., 2021; Kushwaha et al., 2023; Laine et al., 2024; Langer & Landers, 2021; Liu & Tao, 2022; Loureiro et al., 2021; T. Lu & Zhang, 2024; Marabelli et al., 2021; Marjanovic et al., 2022; Marotta, 2022; Martin, 2019; Masiero & Das, 2019; Mirbabaie et al., 2022; Mou & Xu, 2017; Noor et al., 2022; Ochmann et al., 2024; Oleksy et al., 2023; Qin et al., 2023; Rhue, 2024; Russo et al., 2021; Schanke et al., 2021; M. Sharma et al., 2024; Siering, 2022; Sipior et al., 2024; Someh et al., 2019; Tsung-Yu et al., 2024; Wesche & Sonderegger, 2021; Yanit et al., 2023)
Erosion of Meaningful Work	
Loss of Control	(Floridi et al., 2018; Hu et al., 2024; Mirbabaie et al., 2022; Oleksy et al., 2023; Sartori & Theodorou, 2022; Someh et al., 2019)
Overreliance	(Alkhatib, 2021; Bankins et al., 2022; Bauer et al., 2024; Claure et al., 2023; Cowgill et al., 2020; Fleiss et al., 2024; Floridi et al., 2018; Ghasemaghaei & Kordzadeh, 2024a, 2024b; Isaac et al., 2024; Klingbeil et al., 2024; Kushwaha et al., 2023; Laine et al., 2024; Langer & Landers, 2021; Loureiro et al., 2021; T. Lu & Zhang, 2024; Masiero & Das, 2019; Mirbabaie et al., 2022; Ochmann et al., 2024; Porra et al., 2020; Qin et al., 2023; Rhue, 2024; M. Sharma et al., 2024; Siering, 2022; Sipior et al., 2024; Someh et al., 2019; Tsung-Yu et al., 2024; Wang et al., 2021; Wu et al., 2024; Yanit et al., 2023)
Sense of Accountability	(Bankins et al., 2022; Bauer et al., 2024; Bigman et al., 2021; Claure et al., 2023; Feuerriegel et al., 2020; Floridi et al., 2018; Ghasemaghaei & Kordzadeh, 2024a, 2024b; Hoffmann, 2019; Isaac et al., 2024; Kushwaha et al., 2023; Laine et al., 2024; Langer & Landers, 2021; Y. Lu, 2019; Marjanovic et al., 2022; Marotta, 2022; Martin, 2019; Pessach & Shmueli, 2021; Ransbotham et al., 2016; Siering, 2022; Wu et al., 2024; Zhdanov et al., 2022)
Job Displacement	(Aydin et al., 2023; Carter et al., 2020; Claure et al., 2023; Floridi et al., 2018; Heng et al., 2022; Loureiro et al., 2021; Y. Lu, 2019; Marabelli et al., 2021; Mirbabaie et al., 2022; Qin et al., 2023; Ransbotham et al., 2016; Wesche & Sonderegger, 2021; Zhang et al., 2024)
Consolidation of Power	
Global Power Dynamics/Data Colonialism	(Arora et al., 2023; Baudier & de Boissieu, 2025; Buyl & de Bie, 2024; Floridi et al., 2018; Gravett, 2022; Heeks et al., 2024; Heng et al., 2022; Kwet, 2019; Loureiro et al., 2021; Marabelli et al., 2021; Marjanovic et al., 2022; Masiero & Das, 2019; Sartori & Theodorou, 2022; Seoane & Velasco, 2024; Someh et al., 2019; van Berkel et al., 2023; Rieskamp et al. 2023)
Organizational Power Dynamics	(Baudier & de Boissieu, 2025; Floridi et al., 2018; Gravett, 2022; Heng et al., 2022; Kwet, 2019; Marabelli et al., 2021a; Marjanovic et al., 2022; Nedungadi et al., 2024; Seoane & Velasco, 2024; Sipior et al., 2024; Someh et al., 2019; van Berkel et al., 2023)
Algorithmic Control	(Buyl & de Bie, 2024; Floridi et al., 2018; Gravett, 2022; Heeks et al., 2024; Hu et al., 2024; Kwet, 2019; Langer & Landers, 2021; Loureiro et al., 2021; Marabelli et al., 2021; Marjanovic et al., 2022; Ransbotham et al., 2016; Seoane & Velasco, 2024; Someh et al., 2019; Yan :et al., 2024)

Table 4 Categorization of the Literature

Appendix 3: Analysis Process

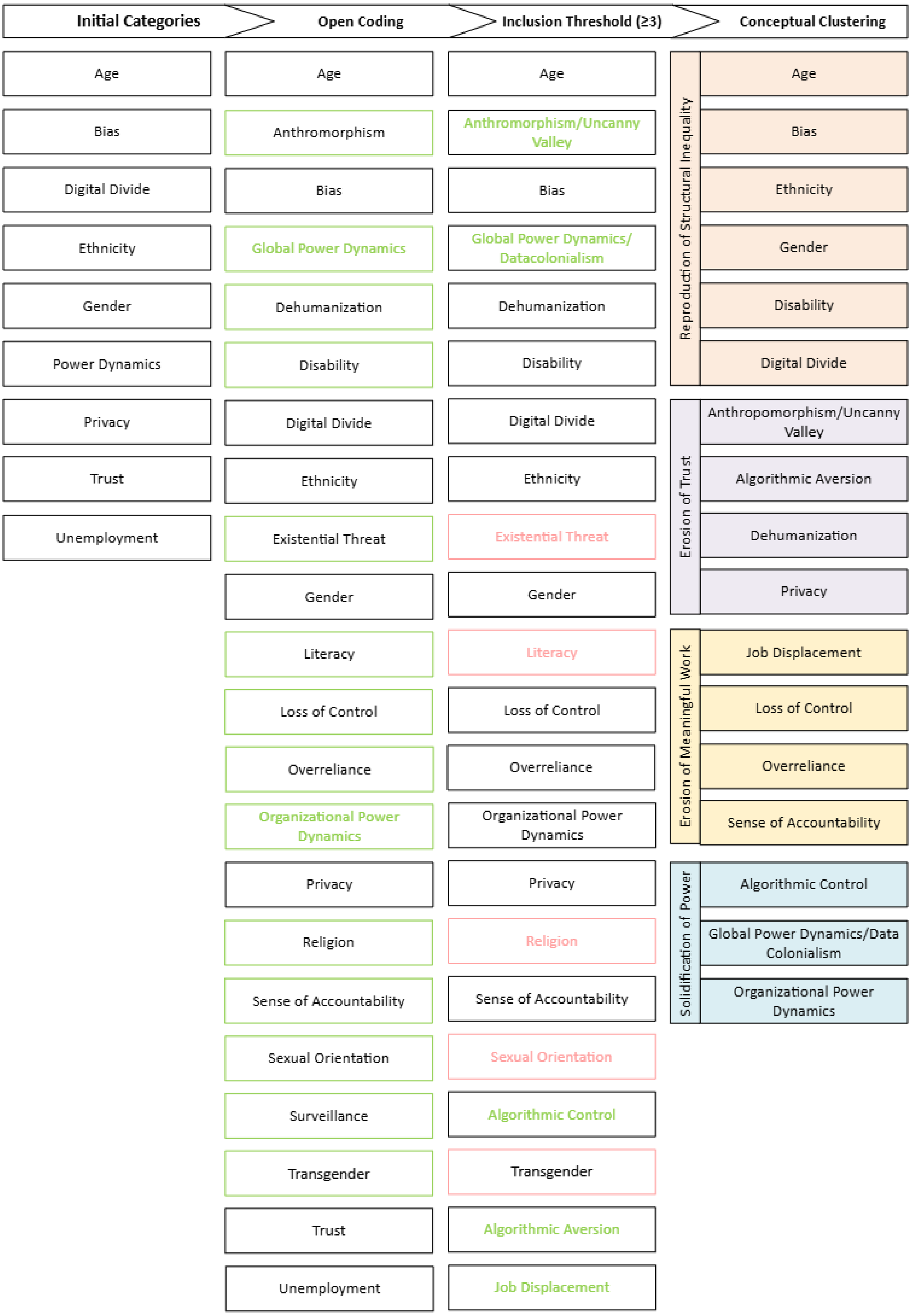

Figure 1 Structured Deductive and Inductive Conceptualisation Process

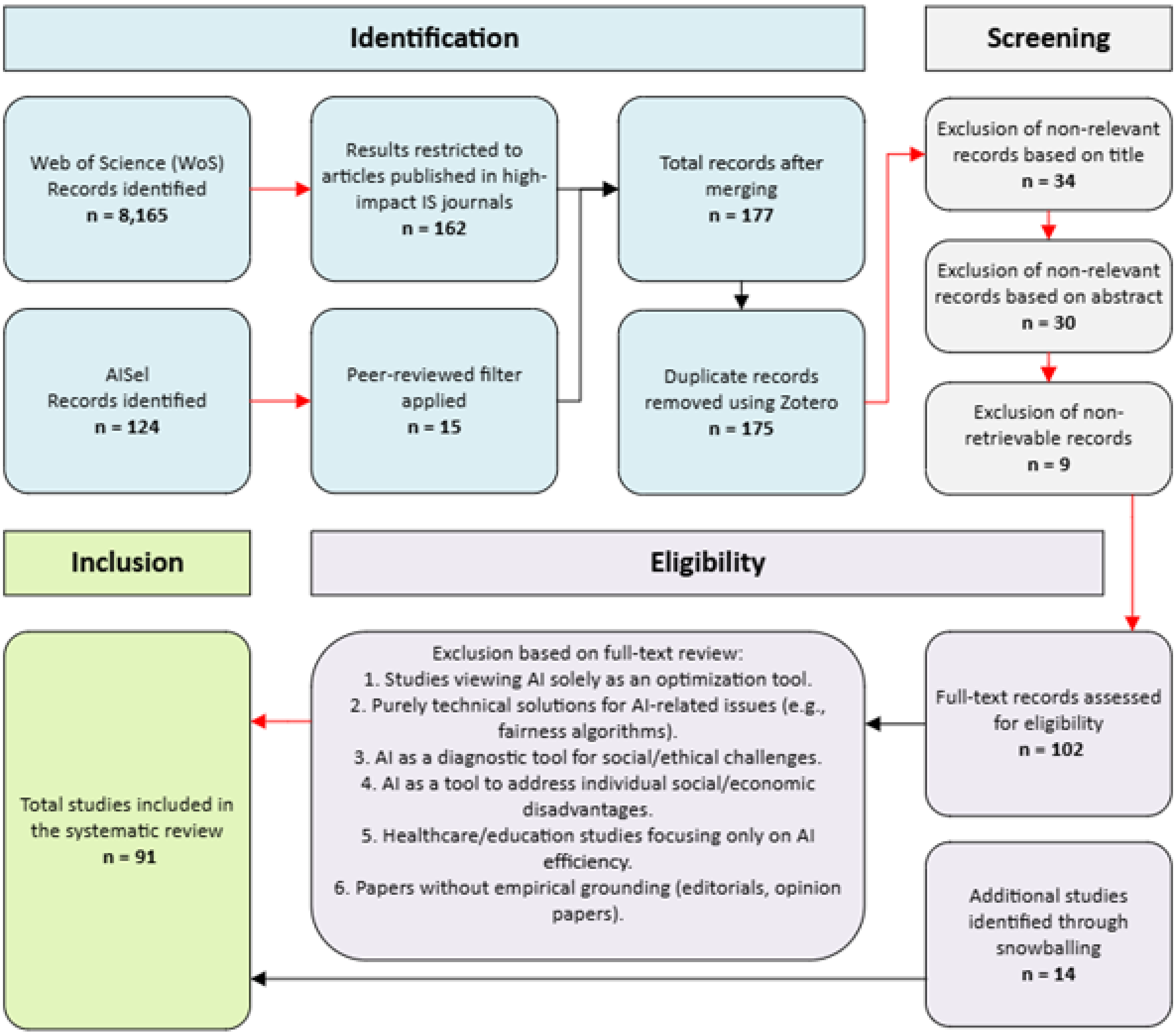

Figure 2 Literature Selection Process for the Systematic Review (based on PRISMA)

Copyright