

HAFIZE NURGUL DURMUS SENYAPAR

AI, organizational governance, and sustainable development: social welfare risks, challenges, and policy implications

Abstract

Research background and purpose: Artificial intelligence (AI) is increasingly embedded in sustainability initiatives, governance arrangements, and digital communication systems, reshaping how information circulates, decisions are made, and social value is distributed. While AI-driven systems enhance efficiency and data-driven coordination, their organizational deployment also intensifies challenges related to algorithmic opacity, misinformation, legitimacy, social inequality, and fragmented governance. Existing research remains fragmented across technical, ethical, and policy-oriented domains and offers limited insight into how organizations simultaneously manage these interconnected impacts. This study aims to develop an organization-centered understanding of AI's societal implications by examining how AI-driven transformations are mediated through organizational decision-making and governance practices across media, culture, social identities, and public policy.

Design/methodology/approach: The study adopts a qualitative, exploratory scoping review methodology. It systematically synthesizes peer-reviewed academic literature and selected international policy frameworks published between 2019 and 2025. Using an inductive thematic synthesis, the review maps recurring patterns and mechanisms through which AI's social implications materialize within organizational contexts, with a focus on conceptual integration rather than causal evaluation.

Findings: The findings show that AI's societal impacts are not direct technological effects but outcomes of organizational governance choices under conditions of fragmented regulation. In digital media, algorithmic ranking, personalization, and moderation restructure information visibility through organizational dependence on opaque platform infrastructures, creating legitimacy and accountability risks. In cultural contexts, AI gains legitimacy through dominant narratives of efficiency and objectivity while reinforcing embedded norms and symbolic authority. Regarding social identities, AI-driven organizational practices reconfigure work roles, agency, and access to resources in socially stratified ways. At the policy level, fragmented AI governance shifts responsibility toward organizations, positioning them as central intermediaries translating regulatory expectations into operational practice.

Value added and limitations: This study contributes by integrating dispersed literatures into an organization-centered framework that links AI, governance, and social impact. It highlights organizational responsibility as a key

determinant of whether AI supports inclusive and sustainable development. The study is limited by its qualitative scoping review design and reliance on secondary sources. Future research should pursue comparative empirical studies and examine organizational decision-making processes in greater depth.

Keywords: *AI governance, organizational responsibility, digital communication, algorithmic bias, sustainable development*

JEL

Classification: O33, Q01, L86, D83

Received: 2025-09-17; **Revised:** 2026-01-22; **Accepted:** 2026-03-31

1. Introduction

Artificial intelligence (AI) has emerged as a transformative force across multiple domains, influencing sustainability, governance, digital communication and broader patterns of social progress. Through advanced data analysis, predictive capabilities and automation, AI offers significant opportunities to address complex societal challenges and to support progress toward the United Nations Sustainable Development Goals (SDGs) (Vinuesa et al., 2020). As a global framework aimed at fostering peace, prosperity and environmental responsibility, the SDGs explicitly recognize the potential of digital and AI-driven technologies to accelerate sustainable development trajectories (Palomares et al., 2021).

At the same time, the integration of AI into decision-making processes introduces profound societal implications that extend beyond technological efficiency. AI-enabled automation has been shown to enhance productivity, optimize resource allocation and stimulate innovation across sectors such as public administration, media and digital governance (Kulal et al., 2024; Suri, 2024). However, the rapid diffusion of AI systems also raises ethical concerns, governance challenges and risks of socio-economic inequality, particularly in relation to transparency, accountability and uneven access to technological benefits (Varoquaux et al., 2025). Importantly, AI's contribution to sustainability is not limited to reductions in energy, water or land use; it increasingly shapes environmental governance practices by informing policy design, regulatory oversight and institutional coordination. Addressing these multifaceted effects requires interdisciplinary approaches that integrate systems thinking, design perspectives and behavioral insights to achieve socially sustainable outcomes (Nishant et al., 2020; Vinuesa et al., 2020).

Although AI research has expanded rapidly across technical, ethical, and policy-oriented disciplines, its societal implications remain unevenly integrated within management and organizational studies. Organizations constitute key sites where

AI-driven transformations are designed, implemented, and governed, acting as intermediaries between technological innovation and broader social outcomes (Shrestha et al., 2019; Radanliev, 2025, 2022). Consequently, understanding AI's societal impacts requires a management-oriented perspective, as organizations operate simultaneously as adopters, governance actors, and strategic decision-makers shaping how AI systems influence media practices, cultural norms, social identities, and public policy processes. This perspective foregrounds organizational decision-making, institutional pressures, legitimacy concerns, and responsibility allocation as central mechanisms through which AI produces social effects (Bankins et al., 2024; Qian et al., 2024; Batool et al., 2025).

Beyond the organizational boundary, AI-driven social impacts increasingly materialize through interconnected domains such as media, culture, social identities and public policy. AI-powered platforms influence public discourse through algorithmic content curation, automated journalism and data-driven recommendation systems, reshaping media trust and democratic communication (Ioscote et al., 2024; Sonni et al., 2024). Simultaneously, AI acts as a cultural producer by transforming creative practices, content generation and symbolic meaning-making, while also redefining professional identities and employment relations through automation and algorithmic management (Cristofaro & Giardino, 2025; Reis & Pinheiro Junior, 2025). In the policy domain, AI both enables evidence-based governance and introduces regulatory challenges related to data privacy, algorithmic bias and fragmented oversight regimes (Fulton et al., 2024; Zhang & Zhang, 2023).

Despite the growing body of literature addressing these dimensions, existing research remains highly fragmented. Studies on AI governance and regulation often privilege legal instruments and ethical principles, analyses of algorithmic bias and inequality focus on distributive consequences, media and communication research examines information flows and public opinion, and management studies concentrate on organizational performance and adoption dynamics (Bircan & Özbilgin, 2025; Birkstedt et al., 2023; Cristofaro & Giardino, 2025; Gehl Sampath, 2021; Pierson et al., 2023; Saeidnia et al., 2025). As a result, existing studies fail to explain how organizations simultaneously manage legitimacy, responsibility, and strategic trade-offs across these domains, despite being central actors in the design, deployment, and governance of AI systems.

From a management and organizational perspective, this fragmentation is particularly consequential. Organizations increasingly operate at the intersection of technological innovation, social expectations and regulatory demands, yet existing studies rarely conceptualize AI's social impacts as interconnected fields that jointly shape organizational legitimacy, responsibility and strategic governance (Beckman et al., 2024; Carter, 2020; Tyagi et al., 2024; Ibáñez & Olmeda, 2022; Kerr et al., 2020). Moreover, much of the current literature is empirically and conceptually skewed toward high-

income and Western contexts, obscuring how AI-related social impacts may vary across different institutional, cultural and regulatory environments (Orr & Davis, 2020; Walter, 2024; Sklavos et al., 2024; Ilcic et al., 2025). These limitations underscore the need for integrative, organization-centered analyses that situate AI's societal implications within broader management and governance frameworks (Af Malmberg & Trondal, 2023; Sifat, 2023; Horneber, 2025; van Noordt et al., 2025).

Within management and organizational scholarship, AI has increasingly been examined in relation to strategy, work organization, governance structures and organizational legitimacy (Camilleri, 2024; Öztürk, 2021). Existing studies largely emphasize efficiency gains, automation and ethical risks, while treating broader social consequences as secondary or indirect outcomes rather than as structurally embedded processes shaping organizational–societal relations (Birkhead et al., 2015; Erman & Furendal, 2024; Rudko et al., 2024). This partial engagement has contributed to the parallel development of research streams on AI governance, media, culture and inequality, limiting their integration within a coherent management-oriented analytical perspective.

The purpose of this study is to develop an organization-centered understanding of the socio-political, ethical, and governance implications of artificial intelligence in the context of sustainable development and digital communication. Rather than examining AI as a purely technical or policy-driven phenomenon, the study focuses on how AI-driven transformations are mediated through organizational decision-making, institutional arrangements, and governance practices across multiple social domains. In doing so, the study advances an integrative analytical perspective that connects AI-driven organizational practices with social outcomes across media, culture, identities, and public policy.

Building on a qualitative exploratory synthesis of interdisciplinary research, the study concentrates on four interconnected fields in which AI's societal impacts are most visible and consequential for organizations: digital media, cultural production, social identities, and public policy. By integrating insights from these domains, the analysis aims to clarify how AI simultaneously reshapes communication ecosystems, cultural norms, identity structures, and governance mechanisms within and around organizations.

Taken together, these considerations point to the need for an organization-centered synthesis that explains how AI-driven social transformations are mediated through managerial and governance practices. Guided by this perspective, the study addresses the following research questions:

RQ1: How do AI-driven organizational practices influence sustainable governance and digital communication ecosystems, particularly in the context of media and information flows?

RQ2: What ethical and policy challenges emerge from organizational uses of AI in relation to algorithmic bias, misinformation, and governance fragmentation?

RQ3: How can organizational strategies and governance frameworks leverage AI technologies to support more equitable and inclusive forms of sustainable development?

By addressing these questions, the study seeks to move beyond descriptive accounts of AI's benefits and risks and toward a more systematic conceptualization of AI-driven social transformation. The analysis contributes to management and organizational scholarship by synthesizing fragmented literatures into an integrated framework that highlights organizations as key mediators of AI's societal effects. In doing so, the study provides a conceptual foundation for future research on AI governance, organizational responsibility, and sustainability-oriented digital transformation.

2. Methods

This study employs a qualitative scoping review to systematically map, organize, and synthesize interdisciplinary research on the social, ethical, and governance implications of AI from a management and organizational perspective. A scoping review is appropriate given the conceptual heterogeneity, fragmented evidence base, and rapid evolution of AI-related scholarship across management, social sciences, media studies, and policy research. Rather than assessing causal effects or intervention outcomes, the objective is analytical and conceptual: to identify dominant patterns, tensions, and gaps that characterize how AI's societal impacts are mediated through organizational decision-making and governance structures.

The review follows an exploratory and interpretive design consistent with established scoping review principles. It focuses on peer-reviewed academic literature and selected international policy frameworks that explicitly address AI in relation to governance, sustainability, digital communication, or organizational contexts. The temporal scope was limited to publications from January 2019 to March 2025 to capture the period in which AI diffusion accelerated alongside intensified regulatory, ethical, and societal debate. This timeframe reflects both the maturation of AI applications and the emergence of global governance initiatives.

The primary academic sources were the Web of Science Core Collection and Scopus databases, selected for their comprehensive coverage of management, social science, and interdisciplinary AI-related research. In addition, analytically relevant policy documents from major international organizations (OECD, UNESCO) and the European Union's AI Act were included to contextualize organizational practices within evolving governance frameworks. These policy texts were treated as conceptual and regulatory reference materials rather than as empirical data sources.

A structured search strategy was applied using database-specific Boolean expressions that combined three conceptual blocks: (1) AI-related terms, (2) social impact domains, and (3) organizational or governance contexts. Searches were conducted across titles, abstracts, and author keywords, and were restricted to English-language, peer-reviewed journal articles. The same conceptual structure was retained across databases, with syntax adapted to platform-specific requirements. For the Web of Science Core Collection, a topic search was conducted using the following query: TS = (("artificial intelligence" OR "AI" OR "algorithmic systems" OR "machine learning" OR "automation") AND (media OR culture OR identity OR identities OR "social identity" OR policy OR governance) AND (organization* OR management OR institution* OR societal)).

For Scopus, the search was applied to titles, abstracts, and keywords using an equivalent Boolean structure: TITLE-ABS-KEY (("artificial intelligence" OR "AI" OR "algorithmic systems" OR "machine learning" OR "automation") AND (media OR culture OR identity OR identities OR "social identity" OR policy OR governance) AND (organization* OR management OR institution* OR societal)). This structured approach ensured conceptual consistency across databases while allowing for database-specific syntax, thereby enhancing transparency and enabling replication of the search process.

Studies were included if they: (i) explicitly addressed AI or algorithmic systems; (ii) examined social, cultural, identity-related, media, or policy implications; and (iii) offered conceptual, empirical, or policy-relevant insights applicable to organizational or societal contexts. Studies were excluded if they focused exclusively on technical or engineering aspects without social analysis, referred to AI only tangentially, or lacked sufficient analytical depth to inform thematic synthesis.

The screening process was conducted in two stages. First, titles and abstracts were reviewed to remove clearly irrelevant records. Second, full texts were assessed against the predefined inclusion and exclusion criteria. Duplicate records were removed prior to screening. This staged process resulted in a purposive and thematically coherent corpus that reflects the interdisciplinary nature of AI-related social research while maintaining a clear management and organizational focus.

For each included study, key analytical information was extracted, including publication context, disciplinary orientation, geographical focus, and core findings related to AI's social implications. A qualitative thematic synthesis was then conducted through iterative coding and comparison. Rather than relying on predefined categories, themes were derived inductively and refined through repeated engagement with the literature and relevant theoretical debates.

This process resulted in four interrelated analytical fields that structure the Results section: (i) AI and digital media, (ii) AI and cultural production, (iii) AI, social identities, and inequality, and (iv) AI, governance, and public policy. These fields

are treated as analytically interconnected rather than isolated domains, reflecting how organizational decision-making mediates AI's societal effects across multiple contexts.

Analytical rigor was ensured through the use of a documented review protocol, transparent inclusion criteria, and iterative theme refinement. The study emphasizes analytical generalization rather than statistical representativeness, aiming to advance conceptual clarity and theory development within management and organizational scholarship. Limitations related to secondary data reliance, Western-centric literature, and the evolving nature of AI research are addressed explicitly in the Discussion section.

3. Results

This section presents a thematic synthesis of the reviewed literature, focusing on recurring patterns, convergences and analytical insights rather than individual study summaries. Drawing on the scoping review process, the findings are organized around four interrelated fields, media, culture, identities and public policy, which together capture the primary ways in which AI reshapes social dynamics across organizational and societal contexts.

3.1. AI, media, and organizational control over information flows

This subsection synthesizes the reviewed literature by focusing on how AI-driven media transformations operate through organizational decision-making and governance structures, rather than summarizing individual empirical cases. The findings highlight recurring organizational mechanisms through which AI reshapes information flows, visibility, and communicative power within digital media ecosystems.

Across the literature, a central and consistent finding is that AI-driven systems fundamentally reconfigure media environments by mediating information visibility, circulation, and credibility. Algorithmic ranking, personalization, and automated moderation systems actively structure public discourse by determining which content is amplified, deprioritized, or removed (Gorwa et al., 2020; Rieder & Skop, 2021; Serttaş et al., 2025). These processes are not neutral technical operations but are embedded in organizational strategies oriented toward engagement optimization, risk management, and commercial performance (Cao et al., 2021; Jarrahi et al., 2021).

The literature further demonstrates that organizational reliance on algorithmic systems increasingly redistributes informational power toward platform-based infrastructures. Media organizations and content producers depend on AI-driven recommendation and distribution mechanisms controlled by dominant platforms, which shapes agenda-setting and narrative control (Sartori & Theodorou, 2022;

Simon, 2022; Voinea, 2025). This dependency constrains organizational autonomy, as decisions regarding visibility and reach are delegated to proprietary systems whose operational logic remains largely opaque (Kinowska & Sienkiewicz, 2023; Zhang et al., 2025).

Opacity emerges as a second dominant analytical pattern. Studies consistently report that algorithmic decision-making processes lack transparency, limiting both organizational oversight and public accountability (Felzmann et al., 2020; Von Eschenbach, 2021). Media organizations often cannot fully explain or contest why specific content is promoted or suppressed, particularly when moderation and ranking are governed by third-party AI systems (Gongane et al., 2022; Gorwa & Veale, 2024). This opacity weakens trust and complicates responsibility allocation in cases of misinformation, over-censorship, or inconsistent enforcement (Gorwa et al., 2020; Ozanne et al., 2022).

Algorithmic bias further intensifies these challenges. Models trained on historical and engagement-driven data tend to reproduce existing social, political, and cultural biases, privileging dominant narratives while marginalizing alternative or minority perspectives (Abokhodair et al., 2024; Willcox, 2025). Recommendation systems optimized for attention and emotional engagement contribute to filter bubbles and ideological polarization, fragmenting democratic debate and public deliberation (Bellina et al., 2023; Park & Park, 2024). These outcomes expose organizations to reputational and legitimacy risks, even when algorithmic decisions are partially externalized to platform providers.

Taken together, the synthesized findings indicate that AI-driven media transformation constitutes a structural reorganization of organizational control over information flows rather than a purely technological upgrade. Media-related challenges such as misinformation, trust erosion, and polarization are closely linked to organizational governance choices concerning AI adoption, oversight, and strategic dependency on platform infrastructures (Lamprou et al., 2025; Simon, 2022). Consequently, AI functions not merely as a media tool but as a force that reshapes organizational decision spaces, redistributes responsibility, and redefines legitimacy in digital communication ecosystems.

3.2. Cultural norms and organizational legitimation of AI

The reviewed literature consistently demonstrates that artificial intelligence is not culturally neutral but operates as a normatively embedded system whose social meaning is produced and stabilized through organizational practices. Across sectors, AI adoption is framed by dominant cultural narratives emphasizing efficiency, objectivity, innovation, and rationality, which organizations mobilize to legitimize algorithmic decision-making internally and externally (Birkstedt et al., 2023; Camilleri, 2024; Gillespie et al., 2023). These narratives position AI not merely as

a technical enhancement but as a symbol of modernity and strategic competence, allowing organizations to justify automation, data-driven control, and restructuring while downplaying social and ethical trade-offs.

A recurrent theme in the literature is that organizational legitimation of AI relies on the alignment between algorithmic systems and prevailing cultural expectations regarding progress and professionalism. AI systems tend to encode dominant values embedded in historical data, institutional routines, and managerial assumptions, thereby reproducing rather than challenging existing social hierarchies (Muldoon & Wu, 2023). Organizations frequently present these outcomes as technically inevitable or culturally appropriate, framing algorithmic outputs as neutral reflections of reality rather than socially constructed judgments. This process contributes to the normalization of algorithmic authority, where AI-driven classifications and recommendations acquire symbolic legitimacy within organizational fields (Öztürk, 2021; Radanliev, 2025).

The literature further highlights cultural ambivalence as a defining condition of AI legitimation. While AI is associated with innovation and competitiveness, it simultaneously evokes concerns related to dehumanization, loss of professional judgment, and erosion of moral responsibility (Bankins et al., 2024; Erman & Furendal, 2024; Rudko et al., 2024). Organizations respond to this tension through discursive strategies that emphasize human oversight, ethical guidelines, and responsibility frameworks, even when decision authority is substantially delegated to algorithmic systems. Such symbolic assurances function as legitimacy buffers, mitigating resistance from employees, users, and regulators without necessarily altering underlying power asymmetries (Bankins et al., 2024; Birkstedt et al., 2023).

Overall, the synthesis indicates that cultural norms play a central role in shaping how AI is interpreted, accepted, and institutionalized within organizations. AI gains legitimacy not through technical performance alone, but through its alignment with culturally embedded notions of rationality, efficiency, and authority. Consequently, organizational adoption of AI should be understood as a cultural process as much as a technological one, where legitimacy is continuously negotiated through narratives, symbols, and governance practices rather than secured by algorithmic capability itself.

3.3. AI, organizational practices, and identity reconfiguration

The literature consistently demonstrates that artificial intelligence reshapes social identities primarily through organizational practices, managerial decision-making, and institutionalized forms of algorithmic governance. Rather than operating as an abstract technological force, AI intervenes in identity formation by reorganizing how individuals are evaluated, categorized, and governed within organizational systems

(Dwivedi et al., 2023). In this respect, identity reconfiguration emerges not merely as an individual-level psychological outcome but as a structurally mediated process embedded in organizational routines, data infrastructures, and performance logics.

Across multiple organizational domains, AI-driven systems increasingly influence access to employment, professional recognition, and material resources. In recruitment and human resource management, algorithmic screening tools rely on historical data to evaluate candidates, often reproducing existing gendered and racialized employment patterns (Wilson & Caliskan, 2024). These systems redefine professional identity by privileging datafied indicators of “fit” and “potential,” thereby shifting agency away from human judgment toward opaque algorithmic criteria. As a result, workers are increasingly positioned as algorithmically legible subjects whose employability and value are continuously assessed through automated metrics rather than contextual expertise or lived experience.

Similar dynamics are observed in public-sector organizations, where AI-supported decision-making systems are deployed in areas such as welfare distribution, taxation, policing, and public administration. Predictive analytics and risk-scoring tools are often justified in terms of efficiency and objectivity; however, empirical evidence indicates that these systems frequently encode historical biases and institutional inequalities (Benjamin, 2019). In welfare administration, automated eligibility assessments have been shown to exclude vulnerable populations without transparent justification or effective appeal mechanisms, leading to identity erosion and a loss of institutional trust (Larsson, 2021). These processes transform citizens’ identities from rights-bearing subjects into risk profiles managed through algorithmic surveillance.

From an organizational perspective, AI-driven identity reconfiguration is unevenly distributed and socially stratified. Highly skilled professionals working in data-intensive or managerial roles may experience identity enhancement, characterized by increased decision authority and symbolic status. In contrast, low-skilled workers and marginalized groups are more likely to experience precarity, deskilling, and exclusion as algorithmic management systems prioritize efficiency and standardization over contextual judgment (Zhang et al., 2025). This asymmetry reinforces existing power relations within organizations and labor markets, positioning AI as a stratifying mechanism rather than a neutral productivity tool.

Beyond employment contexts, AI-driven identity formation is further shaped by organizational control over data, language, and representation. The dominance of English-language datasets and Western-centric training corpora systematically marginalizes non-dominant linguistic and cultural identities, limiting the accuracy and inclusiveness of AI-enabled services (Oloyede, 2025). Organizational decisions regarding dataset selection and model deployment thus have direct implications for whose identities are rendered visible, intelligible, or excluded within digital systems.

Surveillance-oriented organizational practices intensify these identity risks. The integration of facial recognition, biometric tracking, and behavioral analytics into both corporate and public governance infrastructures has raised significant concerns regarding autonomy, consent, and misidentification (Fard et al., 2023; Talay, 2025). Empirical studies demonstrate that these systems disproportionately misidentify women and racialized groups, exposing them to heightened risks of stigmatization and institutional harm. When such technologies are normalized within organizational governance, identity becomes a site of continuous monitoring rather than self-determined expression.

Taken together, the literature indicates that AI-driven identity reconfiguration is best understood as an outcome of organizational design choices rather than as an inevitable consequence of technological advancement. Organizational accountability, transparency mechanisms, and human oversight play a decisive role in determining whether AI systems contribute to empowerment or marginalization (Cheong, 2024; Olatunji Akinrinola et al., 2024). Without deliberate governance interventions, AI risks consolidating asymmetrical power relations and transforming identity into a managerial object governed by algorithmic rationality rather than social justice considerations.

3.4. Organizational responsibility in fragmented AI governance regimes

The synthesized literature converges on a critical finding: contemporary AI governance is characterized by fragmentation across regulatory, institutional, and sectoral boundaries, with organizations positioned at the center of this fragmented landscape. Rather than operating within coherent and harmonized governance regimes, organizations increasingly navigate overlapping, incomplete, or inconsistent regulatory expectations that vary across jurisdictions and policy domains (Veale et al., 2023). This fragmentation shifts governance from a primarily state-centered function toward hybrid arrangements in which organizational decision-making plays a decisive role in shaping the social consequences of AI deployment.

Across the reviewed studies, regulatory divergence emerges as a persistent structural condition rather than a temporary transitional phase. While some contexts—most notably the European Union—have adopted risk-based regulatory frameworks that emphasize accountability, explainability, and proportional oversight, many other regions continue to rely on soft law, voluntary guidelines, or market-driven self-regulation (Laux, 2024). As a result, organizations operating transnationally face asymmetric governance environments, encouraging strategic compliance, selective adaptation, or regulatory arbitrage. The literature indicates that such conditions weaken the enforceability of ethical AI principles and complicate the attribution of responsibility when algorithmic harms occur.

A recurring analytical insight concerns the reallocation of responsibility under fragmented governance regimes. Although governments formally retain regulatory authority, organizations are increasingly expected to internalize governance functions through ethical frameworks, impact assessments, audit mechanisms, and transparency initiatives (De Almeida & Dos Santos Júnior, 2025). This shift effectively positions organizations as intermediaries translating abstract policy principles into operational practices. However, the evidence suggests that organizational self-governance remains uneven in scope and rigor, often constrained by commercial incentives, proprietary technologies, and limited external oversight (Akhai, 2023).

The concentration of AI development and deployment within large technology firms further intensifies these challenges. Dominant platform providers not only control critical data infrastructures and algorithmic architectures but also shape *de facto* global standards through proprietary systems that resist public scrutiny (Mirishli, 2025). The prevalence of opaque “black-box” models restricts the capacity of regulators, civil society, and even adopting organizations to evaluate bias, fairness, or accountability effectively. Consequently, responsibility becomes diffused across complex organizational networks, obscuring causal links between decision-making processes and social outcomes.

Beyond formal regulation, the literature emphasizes the growing normative expectation that organizations contribute proactively to socially responsible AI governance. International organizations such as UNESCO, the OECD, and the G20 promote principles centered on human rights, inclusivity, and sustainability, yet their implementation relies heavily on voluntary organizational commitment (Ko, 2023). In practice, this reliance produces significant variation in how ethical commitments are operationalized, monitored, and enforced. Organizations with greater resources and institutional capacity are more likely to adopt comprehensive governance mechanisms, while others implement symbolic or minimal compliance strategies.

Taken together, the findings suggest that fragmented AI governance regimes do not merely create regulatory uncertainty; they fundamentally reshape organizational responsibility. Organizations are no longer passive rule-takers but active governance actors whose strategic choices influence the distribution of risks, benefits, and accountability associated with AI systems. This positioning underscores a central tension identified across the literature: without clearer alignment between public policy objectives and organizational governance practices, AI risks amplifying inequality, opacity, and institutional failure rather than supporting sustainable and inclusive development (Doyle, 2017).

The synthesized findings across media, culture, identity, and public policy converge on a central insight: the social impacts of AI do not operate in isolation but materialize through organizational decision-making under conditions of fragmented governance. Rather than facing discrete challenges, organizations

confront interconnected pressures related to informational control, normative legitimacy, workforce stratification, and regulatory responsibility. These pressures jointly shape how AI systems are designed, deployed, and justified. Accordingly, effective AI governance cannot be reduced to regulatory compliance alone; it requires organizations to actively coordinate legitimacy management, ethical responsibility, and strategic trade-offs across multiple social domains. Table 1 consolidates these interdependencies by linking each analytical field to the specific organizational tensions and governance responsibilities it generates.

Table 1. **Interconnected fields of AI's social impact and organizational governance implications**

Analytical field	Core social dynamic	Organizational governance tension
Media	Algorithmic control over visibility, ranking, and information circulation	Managing reputation, legitimacy, and accountability under opaque platform dependencies
Culture	Normative ambivalence toward AI as both progress and threat	Securing cultural legitimacy while addressing ethical concerns and resistance to automation
Identity	Uneven reconfiguration of work roles, agency, and professional value	Balancing efficiency-driven automation with equity, inclusion, and workforce stability
Public Policy	Fragmented and uneven AI governance regimes	Translating regulatory expectations into internal governance, self-regulation, and responsibility allocation

Source: own study based on an integrative synthesis of the reviewed literature

3.5. Toward equitable and inclusive organizational AI strategies

The findings indicate that achieving equitable and inclusive outcomes from artificial intelligence depends less on technological capability alone than on organizational strategies, governance choices, and normative commitments. Across sectors, organizations act as the primary intermediaries through which AI systems are designed, procured, implemented, and evaluated. As a result, inequalities associated with AI—such as algorithmic bias, exclusion, and uneven access—are not merely technical failures but outcomes of organizational decision-making and institutional priorities (Tilmes, 2022).

An inclusive organizational AI strategy requires moving beyond efficiency-oriented adoption toward responsibility-oriented governance. Studies emphasize that when organizations prioritize cost reduction, speed, or market dominance, fairness and social inclusion are often treated as secondary considerations (Wang et al., 2021). In contrast, organizations that embed ethical oversight, human review mechanisms, and participatory design principles into AI deployment are better positioned to mitigate discriminatory outcomes and enhance social legitimacy (Panarese et al., 2025). This highlights the importance of governance structures that assign clear responsibility for AI-related harms rather than diffusing accountability across automated systems.

Equity-oriented AI strategies also depend on organizational capacity for transparency and reflexivity. Algorithmic opacity limits the ability of affected individuals and groups to contest decisions, particularly in high-stakes domains such as welfare provision, employment, healthcare, and public administration (Li et al., 2022; Zhong et al., 2025). Research consistently shows that explainability, auditability, and documentation practices are not merely technical features but organizational commitments shaped by leadership priorities and regulatory cultures (Akhai, 2023; Veale et al., 2023). Without such commitments, AI systems risk reinforcing existing power asymmetries while remaining insulated from democratic scrutiny.

Digital literacy and institutional learning emerge as additional pillars of inclusive AI governance. Organizations that invest in AI literacy—not only for technical staff but also for managers, policymakers, and frontline workers—are better equipped to recognize bias, question automated outputs, and intervene when systems produce harmful effects (Strauß, 2021). Conversely, low institutional awareness increases the likelihood of automation bias, where human decision-makers defer uncritically to algorithmic recommendations (Laux & Ruschemeier, 2025).

At a broader level, the literature underscores the need for alignment between organizational practices and emerging international governance frameworks. While global initiatives promoted by organizations such as UNESCO and the OECD emphasize fairness, inclusiveness, and human-centered AI, their effectiveness ultimately depends on how organizations translate these principles into operational rules and everyday practices (Ko, 2023; Lund et al., 2025). Fragmented implementation, symbolic compliance, or ethics-washing can undermine these frameworks, allowing inequitable systems to persist behind formal commitments.

Taken together, these findings suggest that equitable and inclusive AI is not a technological endpoint but an ongoing organizational process. Organizations shape whether AI amplifies social inequalities or contributes to social welfare through choices about governance, accountability, participation, and learning. Ensuring that AI serves inclusive development therefore requires sustained organizational responsibility, continuous evaluation, and a willingness to prioritize social equity alongside innovation and performance.

4. Discussion

This discussion interprets the synthesized findings in relation to the study's research questions, situates the contribution within management and organization studies, and outlines key limitations and directions for future research. In line with the exploratory and integrative scope of the review, the discussion does not introduce new empirical material but analytically connects the results to existing debates on AI, organizational governance, and social impact.

Regarding RQ1 - how AI-driven organizational practices influence sustainable governance and digital communication ecosystems - the findings demonstrate that AI reshapes media and information flows primarily through organizational governance choices rather than through technical features alone. Algorithmic ranking, personalization, and moderation systems affect visibility, credibility, and public discourse insofar as organizations embed these systems into platform strategies, risk management practices, and performance logics. This interpretation aligns with prior work emphasizing the organizational mediation of digital technologies in communication and governance (Medaglia & Tangi, 2022; Pierson et al., 2023). The results extend this literature by showing that media-related challenges such as misinformation, trust erosion, and polarization are closely tied to organizational dependence on opaque platform infrastructures and delegated algorithmic control.

In relation to RQ2, which addresses ethical and policy challenges arising from organizational uses of AI, the synthesis highlights governance fragmentation as a central condition shaping responsibility and accountability. Across the reviewed studies, algorithmic bias, limited explainability, and responsibility diffusion are not isolated ethical failures but structural outcomes of uneven regulatory environments and organizational self-governance practices (Veale et al., 2023). These findings resonate with scholarship on hybrid and soft-law governance of digital technologies (Radu, 2021), while further demonstrating how fragmented regimes shift ethical burden onto organizations without ensuring consistent oversight or enforcement.

Addressing RQ3, the results indicate that equitable and inclusive outcomes from AI depend on organizational strategies that integrate governance, accountability, and learning rather than on technological sophistication alone. Identity-related impacts - particularly in employment, welfare administration, and public services - are shown to emerge through organizational design choices regarding data selection, automation, and oversight. This supports and extends prior research arguing that AI amplifies existing inequalities when organizational incentives prioritize efficiency over inclusion (Xiao, 2025). Conversely, the findings suggest that organizations with stronger governance capacities, transparency mechanisms, and human oversight are better positioned to mitigate exclusionary effects and align AI deployment with broader social objectives.

The primary contribution of this study lies in its organization-centered integration of AI's social implications across media, culture, identity, and public policy. While existing research often examines these domains in isolation, the synthesis demonstrates that their effects are interconnected through recurring organizational mechanisms, including algorithmic mediation, normative legitimation, and responsibility allocation. By foregrounding organizations as active intermediaries rather than passive adopters of AI, the study advances management and organization scholarship that seeks to move beyond performance-oriented accounts of digital transformation.

Conceptually, the findings contribute to ongoing debates on digital governance and organizational responsibility by framing AI's social impacts as contingent on institutional context, organizational capability, and power asymmetries. This process-oriented perspective complements prior work on digital transformation and technology adoption by explicitly linking AI deployment to legitimacy management, workforce stratification, and governance alignment (Canbul Yaroğlu, 2024; Humberd & Latham, 2025). Rather than proposing universal effects of AI, the study clarifies the conditions under which similar technologies generate divergent outcomes across organizational and regulatory settings.

Several limitations should be acknowledged. First, the study is based on an exploratory scoping review of secondary sources, which limits the ability to draw context-specific or causal inferences. While this approach is appropriate for synthesizing fragmented literatures, it does not capture the lived experiences of organizational actors or the micro-level dynamics of AI implementation. Second, the global scope of the review may obscure regional variations in governance capacity, cultural acceptance, and institutional design, particularly in low- and middle-income contexts. Third, although the analysis addresses social and governance dimensions of AI, it does not provide quantitative assessments of environmental or economic impacts associated with large-scale AI deployment.

Future research could address these limitations by conducting comparative empirical studies across organizational and national contexts, with particular attention to under-researched regions and sectors. Mixed-method designs combining qualitative insights with quantitative measures would be valuable for examining how organizational decision-making processes shape AI outcomes in practice. Further work is also needed to explore how emerging regulatory frameworks are translated into internal governance mechanisms and how organizational learning and digital literacy influence the long-term sustainability and inclusiveness of AI-driven transformation.

Overall, the discussion underscores that AI's societal consequences are neither technologically predetermined nor socially uniform. They are produced through organizational choices made under conditions of fragmented governance and unequal power. Understanding AI as an organizationally mediated process is therefore essential for both theory development and the design of responsible, equitable, and sustainable AI strategies.

5. Conclusions

This study synthesized interdisciplinary research on AI's social implications through an organization-centered lens, focusing on how AI-related effects are mediated by managerial choices, governance arrangements, and institutional environments. Across the reviewed literature, three integrated conclusions address the research questions.

Regarding RQ1, AI reshapes digital communication ecosystems primarily through organizationally embedded algorithmic governance. Systems for ranking, personalization, and moderation reorganize information visibility and credibility, but their social consequences are contingent on how organizations adopt, outsource, and oversee these infrastructures. The evidence consistently indicates that organizational dependency on opaque platform logics amplifies reputational exposure and complicates accountability in cases of misinformation, inconsistent enforcement, and trust erosion.

Regarding RQ2, the literature converges on governance fragmentation as a structural condition that diffuses responsibility and weakens enforceability of ethical AI principles. Divergent regulatory models, uneven institutional capacity, and the prevalence of proprietary “black-box” systems encourage selective compliance and complicate harm attribution across organizational networks. In this setting, organizations become de facto governance actors, yet self-governance remains uneven and often constrained by commercial incentives and limited external scrutiny.

Regarding RQ3, equity and inclusion emerge less as technical outputs of AI capability and more as outcomes of organizational strategy, governance commitment, and institutional learning. AI-driven identity reconfiguration in employment and public services is consistently shown to be socially stratifying when efficiency-oriented automation displaces contextual judgment, particularly for marginalized groups. The evidence suggests that transparency, auditability, meaningful human oversight, and AI literacy are necessary organizational conditions for mitigating exclusionary effects and sustaining legitimacy.

The study's main contribution is conceptual and integrative. It advances management and organization studies by showing that AI's social implications across media, culture, identity, and public policy are not independent “impact areas” but mutually reinforcing domains that materialize through organizational decision-making under fragmented governance. By foregrounding legitimacy management, responsibility allocation, and strategic trade-offs as cross-cutting mechanisms, the study addresses the literature gap identified in the Introduction: why existing research struggles to explain how organizations simultaneously manage competing pressures across these domains despite being central intermediaries of AI deployment. In doing so, it complements sustainability-oriented AI debates by linking SDG-relevant ambitions to organizational governance

realities rather than treating sustainability effects as detached from institutions and managerial practice.

For organizational leaders and managers, the findings imply that AI adoption should be governed as a legitimacy- and responsibility-intensive transformation, not an operational optimization project. Decisions about procurement, outsourcing to platform infrastructures, model transparency, and human oversight directly shape trust, reputational risk, and the organization's exposure to accountability disputes.

For policymakers and regulators, the synthesis highlights that rule design alone is insufficient when implementation relies heavily on organizational self-governance in uneven regulatory landscapes. Effective governance requires enforceable alignment mechanisms that translate high-level principles into auditable organizational practices, reducing incentives for regulatory arbitrage and responsibility diffusion.

For platform firms and AI providers, the evidence underscores that opacity is not merely a technical limitation but a governance choice with downstream social and institutional consequences. Increasing explainability, documentation, and contestability is central to credible accountability in high-stakes contexts, particularly where third-party systems shape information flows and eligibility decisions.

For civil society, professional bodies, and educators, the findings support investment in AI literacy as an institutional capability—not only for technical staff, but for managers and frontline decision-makers—so that automated outputs can be questioned and corrected rather than deferred to by default.

This study has limitations consistent with its scoping-review design. First, the synthesis prioritizes conceptual mapping over causal inference; it identifies patterns and mechanisms but does not estimate effect sizes or test causality. Second, the corpus is restricted to English-language peer-reviewed journal articles and selected policy frameworks, which may underrepresent locally grounded scholarship and non-English debates, particularly from low- and middle-income contexts highlighted as under-examined in the Introduction. Third, AI is treated as a broad category to preserve analytical coherence; this may mask variation across distinct AI applications (e.g., generative systems, algorithmic management, decision-support in welfare administration).

Future research can build on these constraints by (i) conducting comparative empirical studies of organizational AI governance across regulatory regimes and cultural settings, (ii) disaggregating AI applications to test which configurations most strongly shape legitimacy, accountability, and inclusion outcomes, and (iii) examining the internal organizational processes through which AI governance is negotiated (e.g., conflict between compliance, performance, and ethics functions), including longitudinal designs that capture change over time.

The reviewed evidence supports a clear practical conclusion: socially sustainable AI is not achieved by technological performance alone. It requires organizations to

operationalize responsibility through governance structures that make AI systems transparent enough to be accountable, inclusive enough to avoid systematic exclusion, and aligned enough with public policy objectives to prevent fragmented oversight from becoming institutionalized failure. International principle-setting efforts remain important, but their societal value depends on how organizations convert them into enforceable routines, auditable controls, and learning-oriented practices in everyday decision-making.

Acknowledgments

The author would like to acknowledge that an earlier version of this study was presented as an oral presentation at the *NEWPATH New Paradigms in Communication Technologies and Humanity Symposium: Artificial Intelligence – Technological, Social and Cultural Ties*, held on 6–7 March 2025 in İstanbul, Türkiye.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used *Grammarly* and *Quillbot* in order to check language accuracy, clarity and readability. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

References

- Abokhodair, N., Skop, Y., Rüller, S., Aal, K., & Elmimouni, H. (2024). Opaque algorithms, transparent biases: Automated content moderation during the Sheikh Jarrah crisis. *First Monday*, 29(4). <https://doi.org/10.5210/fm.v29i4.13620>
- Af Malmborg, F., & Trondal, J. (2023). Discursive framing and organizational venues: Mechanisms of artificial intelligence policy adoption. *International Review of Administrative Sciences*, 89(1), 39–58. <https://doi.org/10.1177/00208523211007533>
- Akhai, S. (2023). From black boxes to transparent machines: The quest for explainable AI. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4390887>
- Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior*, 45(2), 159–182. <https://doi.org/10.1002/job.2735>
- Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. *AI and Ethics*, 5(3), 3265–3279. <https://doi.org/10.1007/s43681-024-00653-w>
- Beckman, L., Hultin Rosenberg, J., & Jebari, K. (2024). Artificial intelligence and democratic legitimacy: The problem of publicity in public authority. *AI & Society*, 39(3), 975–984. <https://doi.org/10.1007/s00146-022-01493-0>

- Bellina, A., Castellano, C., Pineau, P., Iannelli, G., & De Marzo, G. (2023). Effect of collaborative-filtering-based recommendation algorithms on opinion polarization. *Physical Review E*, 108(5), 054304. <https://doi.org/10.1103/PhysRevE.108.054304>
- Benjamin, R. (2019). Assessing risk, automating racism. *Science*, 366(6464), 421–422. <https://doi.org/10.1126/science.aaz3873>
- Bircan, T., & Özbilgin, M. F. (2025). Unmasking inequalities of the code: Disentangling the nexus of AI and inequality. *Technological Forecasting and Social Change*, 211, 123925. <https://doi.org/10.1016/j.techfore.2024.123925>
- Birkhead, G. S., Klompas, M., & Shah, N. R. (2015). Uses of electronic health records for public health surveillance to advance public health. *Annual Review of Public Health*, 36(1), 345–359. <https://doi.org/10.1146/annurev-publhealth-031914-122747>
- Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. *Internet Research*, 33(7), 133–167. <https://doi.org/10.1108/INTR-01-2022-0042>
- Camilleri, M. A. (2024). Artificial intelligence governance: Ethical considerations and implications for social responsibility. *Expert Systems*, 41(7), e13406. <https://doi.org/10.1111/exsy.13406>
- Canbul Yaroğlu, A. (2024). Who's in the mirror: Shaping organizational identity through artificial intelligence and symbolic interactionism. *Kybernetes*, 55(2), 1255–1277. <https://doi.org/10.1108/K-09-2024-2379>
- Cao, B., Zhao, J., Lv, Z., & Yang, P. (2021). Diversified personalized recommendation optimization based on mobile data. *IEEE Transactions on Intelligent Transportation Systems*, 22(4), 2133–2139. <https://doi.org/10.1109/TITS.2020.3040909>
- Carter, D. (2020). Regulation and ethics in artificial intelligence and machine learning technologies: Where are we now? Who is responsible? Can the information professional play a role? *Business Information Review*, 37(2), 60–68. <https://doi.org/10.1177/0266382120923962>
- Cheong, B. C. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, 6, 1421273. <https://doi.org/10.3389/fhumd.2024.1421273>
- Cristofaro, M., & Giardino, P. L. (2025). Surfing the AI waves: The historical evolution of artificial intelligence in management and organizational studies and practices. *Journal of Management History*. Ahead-of-print. <https://doi.org/10.1108/JMH-01-2025-0002>
- De Almeida, P. G. R., & Dos Santos Júnior, C. D. (2025). Artificial intelligence governance: Understanding how public organizations implement it. *Government Information Quarterly*, 42(1), 102003. <https://doi.org/10.1016/j.giq.2024.102003>
- Doyle, T. (2017). Weapons of math destruction: How big data increases inequality and threatens democracy (Cathy O'Neil). *The Information Society*, 33(5), 301–302. <https://doi.org/10.1080/01972243.2017.1354593>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Erman, E., & Furendal, M. (2024). Artificial intelligence and the political legitimacy of global governance. *Political Studies*, 72(2), 421–441. <https://doi.org/10.1177/00323217221126665>

- Fard, N. E., Selmic, R. R., & Khorasani, K. (2023). Public policy challenges, regulations, oversight, technical, and ethical considerations for autonomous systems: A survey. *IEEE Technology and Society Magazine*, 42(1), 45–53. <https://doi.org/10.1109/MTS.2023.3241315>
- Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2020). Towards transparency by design for artificial intelligence. *Science and Engineering Ethics*, 26(6), 3333–3361. <https://doi.org/10.1007/s11948-020-00276-4>
- Fulton, R., Fulton, D., Hayes, N., & Kaplan, S. (2024). The transformation risk-benefit model of artificial intelligence: Balancing risks and benefits through practical solutions and use cases. *International Journal of Artificial Intelligence & Applications*, 15(2), 1–22. <https://doi.org/10.5121/ijaia.2024.15201>
- Gehl Sampath, P. (2021). Governing artificial intelligence in an age of inequality. *Global Policy*, 12(S6), 21–31. <https://doi.org/10.1111/1758-5899.12940>
- Gillespie, N., Lockey, S., Curtis, C., Pool, J., & Ali Akbari. (2023). *Trust in artificial intelligence: A global study* (Technical report). The University of Queensland; KPMG Australia. <https://espace.library.uq.edu.au/view/UQ:00d3c94>
- Gongane, V. U., Munot, M. V., & Anuse, A. D. (2022). Detection and moderation of detrimental content on social media platforms: Current status and future directions. *Social Network Analysis and Mining*, 12(1), 129. <https://doi.org/10.1007/s13278-022-00951-3>
- Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1). <https://doi.org/10.1177/2053951719897945>
- Gorwa, R., & Veale, M. (2024). Moderating model marketplaces: Platform governance puzzles for AI intermediaries. *Law, Innovation and Technology*, 16(2), 341–391. <https://doi.org/10.1080/17579961.2024.2388914>
- Horneber, D. (2025). Understanding the implementation of responsible artificial intelligence in organizations: A neo-institutional theory perspective. *Communications of the Association for Information Systems*, 57(1), 185–218. <https://doi.org/10.17705/1CAIS.05708>
- Humberd, B. K., & Latham, S. F. (2025). When AI becomes an agent of the firm: Examining the evolution of AI in organizations through an agency theory lens. *Journal of Management Studies*, 63, 668–694. <https://doi.org/10.1111/joms.13274>
- Ibáñez, J. C., & Olmeda, M. V. (2022). Operationalising AI ethics: How are companies bridging the gap between practice and principles? An exploratory study. *AI & Society*, 37(4), 1663–1687. <https://doi.org/10.1007/s00146-021-01267-0>
- Ilcic, A., Fuentes, M., & Lawler, D. (2025). Artificial intelligence, complexity, and systemic resilience in global governance. *Frontiers in Artificial Intelligence*, 8, 1562095. <https://doi.org/10.3389/frai.2025.1562095>
- Ioscote, F., Gonçalves, A., & Quadros, C. (2024). Artificial intelligence in journalism: A ten-year retrospective of scientific articles (2014–2023). *Journalism and Media*, 5(3), 873–891. <https://doi.org/10.3390/journalmedia5030056>
- Jarrahi, M. H., Newlands, G., Lee, M. K., Wolf, C. T., Kinder, E., & Sutherland, W. (2021). Algorithmic management in a work context. *Big Data & Society*, 8(2), 20539517211020332. <https://doi.org/10.1177/20539517211020332>
- Kerr, A., Barry, M., & Kelleher, J. D. (2020). Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance. *Big Data & Society*, 7(1), 2053951720915939. <https://doi.org/10.1177/2053951720915939>

- Kinowska, H., & Sienkiewicz, Ł. J. (2023). Influence of algorithmic management practices on workplace well-being: Evidence from European organisations. *Information Technology & People*, 36(8), 21–42. <https://doi.org/10.1108/ITP-02-2022-0079>
- Ko, B.-M. (2023). Analysis of international regulations on artificial intelligence (AI) ethics: A comparative approach. *Asia-Pacific Journal of Business & Commerce*, 15(3), 201–225. <https://doi.org/10.35183/ajbc.2023.11.15.3.201>
- Kulal, A., Rahiman, H. U., Suvarna, H., Abhishek, N., & Dinesh, S. (2024). Enhancing public service delivery efficiency: Exploring the impact of AI. *Journal of Open Innovation: Technology, Market, and Complexity*, 10(3), 100329. <https://doi.org/10.1016/j.joitmc.2024.100329>
- Lamprou, S., Dekoulou, P. (Evi), & Kalliris, G. (2025). The critical impact and socio-ethical implications of AI on content generation practices in media organizations. *Societies*, 15(8), 214. <https://doi.org/10.3390/soc15080214>
- Larsson, K. K. (2021). Digitization or equality: When government automation covers some, but not all citizens. *Government Information Quarterly*, 38(1), 101547. <https://doi.org/10.1016/j.giq.2020.101547>
- Laux, J. (2024). Institutionalised distrust and human oversight of artificial intelligence: Towards a democratic design of AI governance under the European Union AI Act. *AI & Society*, 39(6), 2853–2866. <https://doi.org/10.1007/s00146-023-01777-z>
- Laux, J., & Ruschemeier, H. (2025). Automation bias in the AI Act: On the legal implications of attempting to de-bias human oversight of AI. *European Journal of Risk Regulation*, 16(4), 1519–1534. <https://doi.org/10.1017/err.2025.10033>
- Li, F., Ruijs, N., & Lu, Y. (2022). Ethics & AI: A systematic review on ethical concerns and related strategies for designing with AI in healthcare. *AI*, 4(1), 28–53. <https://doi.org/10.3390/ai4010003>
- Lund, B., Orhan, Z., Mannuru, N. R., Bevara, R. V. K., Porter, B., Vinaih, M. K., & Bhaskara, P. (2025). Standards, frameworks, and legislation for artificial intelligence (AI) transparency. *AI and Ethics*, 5(4), 3639–3655. <https://doi.org/10.1007/s43681-025-00661-4>
- Medaglia, R., & Tangi, L. (2022). The adoption of artificial intelligence in the public sector in Europe: Drivers, features, and impacts. In P. Parycek, T. Kriplean, & T. Escher (Eds.), *ICEGOV 2022: Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance* (pp. 10–18). Association for Computing Machinery. <https://doi.org/10.1145/3560107.3560110>
- Mirishli, S. (2025). The role of legal frameworks in shaping ethical artificial intelligence use in corporate governance. *arXiv*. <https://doi.org/10.48550/arXiv.2503.14540>
- Muldoon, J., & Wu, B. A. (2023). Artificial intelligence in the colonial matrix of power. *Philosophy & Technology*, 36(4), 80. <https://doi.org/10.1007/s13347-023-00687-8>
- Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management*, 53, 102104. <https://doi.org/10.1016/j.ijinfomgt.2020.102104>
- Olatunji Akinrinola, C., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Advanced Research and Reviews*, 18(3), 50–58. <https://doi.org/10.30574/gscarr.2024.18.3.0088>
- Oloyede, J. (2025). Leveraging artificial intelligence to foster tolerance and coexistence in multicultural societies: A framework for inclusive policy and social integration. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5318927>

- Orr, W., & Davis, J. L. (2020). Attributions of ethical responsibility by artificial intelligence practitioners. *Information, Communication & Society*, 23(5), 719–735. <https://doi.org/10.1080/1369118X.2020.1713842>
- Ozanne, M., Bhandari, A., Bazarova, N. N., & DiFranzo, D. (2022). Shall AI moderators be made visible? Perception of accountability and trust in moderation systems on social media platforms. *Big Data & Society*, 9(2), 20539517221115666. <https://doi.org/10.1177/20539517221115666>
- Öztürk, D. (2021). What does artificial intelligence mean for organizations? A systematic review of organization studies research and a way forward. In S. Bozkuş Kahyaoglu (Ed.), *The impact of artificial intelligence on governance, economics and finance* (Vol. 1, pp. 265–289). Springer Nature. https://doi.org/10.1007/978-981-33-6811-8_14
- Palomares, I., Martínez-Cámara, E., Montes, R., García-Moral, P., Chiachio, M., Chiachio, J., Alonso, S., Melero, F. J., Molina, D., Fernández, B., Moral, C., Marchena, R., De Vargas, J. P., & Herrera, F. (2021). A panoramic view and SWOT analysis of artificial intelligence for achieving the sustainable development goals by 2030: Progress and prospects. *Applied Intelligence*, 51(9), 6497–6527. <https://doi.org/10.1007/s10489-021-02264-y>
- Park, H. W., & Park, S. (2024). The filter bubble generated by artificial intelligence algorithms and the network dynamics of collective polarization on YouTube: The case of South Korea. *Asian Journal of Communication*, 34(2), 195–212. <https://doi.org/10.1080/01292986.2024.2315584>
- Panarese, P., Grasso, M. M., & Solinas, C. (2025). Algorithmic bias, fairness, and inclusivity: A multilevel framework for justice-oriented AI. *AI & Society*. <https://doi.org/10.1007/s00146-025-02451-2>
- Pierson, J., Kerr, A., Robinson, S. C., Fanni, R., Steinkogler, V. E., Milan, S., & Zampedri, G. (2023). Governing artificial intelligence in the media and communications sector. *Internet Policy Review*, 12(1), 1–28. <https://doi.org/10.14763/2023.1.1683>
- Qian, Y., Siau, K. L., & Nah, F. F. (2024). Societal impacts of artificial intelligence: Ethical, legal, and governance issues. *Societal Impacts*, 3, 100040. <https://doi.org/10.1016/j.socimp.2024.100040>
- Radanliev, P. (2025). Artificial intelligence: Reflecting on the past and looking towards the next paradigm shift. *Journal of Experimental & Theoretical Artificial Intelligence*, 37(7), 1045–1062. <https://doi.org/10.1080/0952813X.2024.2323042>
- Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, 40(2), 178–193. <https://doi.org/10.1080/14494035.2021.1929728>
- Reis, J. F., & Pinheiro Junior, L. P. (2025). Institutional theory (IT) and diffusion of innovation (DOI): A theoretical approach on artificial intelligence (AI). *BAR - Brazilian Administration Review*, 22, e250060. <https://doi.org/10.1590/1807-7692bar2025250060>
- Rieder, B., & Skop, Y. (2021). The fabrics of machine moderation: Studying the technical, normative, and organizational structure of Perspective API. *Big Data & Society*, 8(2), 20539517211046181. <https://doi.org/10.1177/20539517211046181>
- Rudko, I., Bashirpour Bonab, A., Fedele, M., & Formisano, A. V. (2024). New institutional theory and AI: Toward rethinking of artificial intelligence in organizations. *Journal of Management History*, 31(2), 261–284. <https://doi.org/10.1108/JMH-09-2023-0097>
- Saeidnia, H. R., Hosseini, E., Lund, B., Tehrani, M. A., Zaker, S., & Molaei, S. (2025). Artificial intelligence in the battle against disinformation and misinformation: A systematic review of challenges and approaches. *Knowledge and Information Systems*, 67(4), 3139–3158. <https://doi.org/10.1007/s10115-024-02337-7>

- 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>
- Serttaş, A., Gürkan, H., & Dere, G. (2025). Synthetic social alienation: The role of algorithm-driven content in shaping digital discourse and user perspectives. *Journalism and Media*, 6(3), 149. <https://doi.org/10.3390/journalmedia6030149>
- Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, 61(4), 66–83. <https://doi.org/10.1177/0008125619862257>
- Sifat, I. (2023). Artificial intelligence as an institution: Impacts, integration strategies, and policy directions (SSRN Scholarly Paper No. 4545403). *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4545403>
- Simon, F. M. (2022). Uneasy bedfellows: AI in the news, platform companies and the issue of journalistic autonomy. *Digital Journalism*, 10(10), 1832–1854. <https://doi.org/10.1080/21670811.2022.2063150>
- Sklavos, G., Theodossiou, G., Papanikolaou, Z., Karelakis, C., & Ragazou, K. (2024). Environmental, social, and governance-based artificial intelligence governance: Digitalizing firms' leadership and human resources management. *Sustainability*, 16(16), 7154. <https://doi.org/10.3390/su16167154>
- Sonni, A. F., Hafied, H., Irwanto, I., & Latuheru, R. (2024). Digital newsroom transformation: A systematic review of the impact of artificial intelligence on journalistic practices, news narratives, and ethical challenges. *Journalism and Media*, 5(4), 1554–1570. <https://doi.org/10.3390/journalmedia5040097>
- Strauß, S. (2021). “Don't let me be misunderstood”: Critical AI literacy for the constructive use of AI technology. *TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis*, 30(3), 44–49. <https://doi.org/10.14512/tatup.30.3.44>
- Suri, S. (2024). Defining our future with generative AI. *Nature Computational Science*, 4(9), 641–643. <https://doi.org/10.1038/s43588-024-00694-5>
- Talay, Ö. (2025). How surveillance shapes migrant information practices and social integration in Turkey. *Information Development*. Advance online publication. <https://doi.org/10.1177/02666669251373044>
- Tyagi, L., Gupta, A., & Singh Sisodia, V. (2024). Revolutionizing Industries: AI-Driven Case Studies and Success Stories in Real-World Applications and Innovations. In *Proceedings of the 2024 1st International Conference on Sustainable Computing and Integrated Communication in Changing Landscape of AI (ICSCAI)* (pp. 1–9). IEEE. <https://doi.org/10.1109/ICSCAI61790.2024.10866349>
- Tilmes, N. (2022). Disability, fairness, and algorithmic bias in AI recruitment. *Ethics and Information Technology*, 24(2), 21. <https://doi.org/10.1007/s10676-022-09633-2>
- van Noordt, C., Medaglia, R., & Tangi, L. (2025). Policy initiatives for artificial intelligence-enabled government: An analysis of national strategies in Europe. *Public Policy and Administration*, 40(2), 215–253. <https://doi.org/10.1177/09520767231198411>
- Varoquaux, G., Luccioni, S., & Whittaker, M. (2025). Hype, sustainability, and the price of the bigger-is-better paradigm in AI. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency* (pp. 61–75). <https://doi.org/10.1145/3715275.3732006>
- Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, 19(1), 255–275. <https://doi.org/10.1146/annurev-lawsocsci-020223-040749>

- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the sustainable development goals. *Nature Communications*, *11*(1), 233. <https://doi.org/10.1038/s41467-019-14108-y>
- Voinea, D. V. (2025). Reconceptualizing gatekeeping in the age of artificial intelligence: A theoretical exploration of artificial intelligence-driven news curation and automated journalism. *Journalism and Media*, *6*(2), 68. <https://doi.org/10.3390/journalmedia6020068>
- Von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology*, *34*(4), 1607–1622. <https://doi.org/10.1007/s13347-021-00477-0>
- Walter, Y. (2024). Managing the race to the moon: Global policy and governance in artificial intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. *Discover Artificial Intelligence*, *4*(1), 14. <https://doi.org/10.1007/s44163-024-00109-4>
- Wang, Y., Su, M., Shen, L., & Tang, R. (2021). Decision-making of closed-loop supply chain under corporate social responsibility and fairness concerns. *Journal of Cleaner Production*, *284*, 125373. <https://doi.org/10.1016/j.jclepro.2020.125373>
- Willcox, M. (2025). Algorithmic agency and “fighting back” against discriminatory Instagram content moderation: #IWantToSeeNyome. *Frontiers in Communication*, *9*, 1385869. <https://doi.org/10.3389/fcomm.2024.1385869>
- Wilson, K., & Caliskan, A. (2024). Gender, race, and intersectional bias in resume screening via language model retrieval. *arXiv*. <https://doi.org/10.48550/arXiv.2407.20371>
- Xiao, J. (2025). Secondary bounded rationality: A theory of how algorithms reproduce structural inequality in AI hiring [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2507.09233>
- Zhang, J., & Zhang, Z. (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC Medical Informatics and Decision Making*, *23*(1), 7. <https://doi.org/10.1186/s12911-023-02103-9>
- Zhang, M. M., Cooke, F. L., Ahlstrom, D., & McNeil, N. (2025). The rise of algorithmic management and implications for work and organisations. *New Technology, Work and Employment*, *40*(3), 659–671. <https://doi.org/10.1111/ntwe.12343>
- Zhong, J., Shetty, A., Jia, C., Lin, X., & Naseem, U. (2025). Pluralistic alignment for healthcare: A role-driven framework. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing* (pp. 31308–31331). <https://doi.org/10.18653/v1/2025.emnlp-main.1596>