

Human-Centric Governance for Digital Trust: A Systematic Literature Review of Responsible AI in Data-Driven Societies

Rajat Verma¹, Shashi Kant Gupta²

^{1,2}Lincoln University College, Petaling Jaya, Selangor-47301, Malaysia

¹pdf.rajatverma@lincoln.edu.my, ²shashigupta@lincoln.edu.my

Abstract: With the increasing use of Artificial Intelligence (AI) in Data-Driven Societies, a new set of governance issues has arisen, such as Privacy, Transparency, Accountability, Ethical Responsibility, Human Monitoring and Trust. While many of these issues are covered in Responsible AI research, the information is spread across principles, technologies, and trust issues. This makes it hard to comprehend how Human-Centric Governance can construct Digital Trust. This study gives a snapshot of recent research on Responsible AI to appreciate its links to Human-Focused Governance and Digital Trust. According to the review, Privacy Protection, Transparency, Accountability, Ethical AI Governance and Human Monitoring are important and interrelated factors that influence trust in an AI environment. The findings highlight the need to more closely link these governance areas to Digital Trust. This study integrates existing studies to provide a Human-Centered approach to governance with Digital Trust at its core, defining Responsible AI. The findings are expected to inform researchers, policymakers, organizations, and technology practitioners in designing governance approaches to promote responsible, accountable, transparent, and trustworthy AI systems.

Keywords: Digital Trust; Human-Centric Governance; Responsible AI; Systematic Literature Review; Trustworthy AI.

Introduction

AI (Artificial Intelligence) is rapidly evolving and becoming more data-driven in societies. Nowadays, AI is employed in various sectors, such as business, digital platforms, public services, decision-making, and more, to showcase its capabilities. This presents numerous opportunities for growth and productivity as well as governance challenges [1,2].

Responsible AI is crucial for humanizing AI systems to ensure they behave responsibly, transparently, in line with human values and respect privacy. Recent research papers emphasize, the importance of strong governance frameworks, relationships and processes for Responsible AI [1]. AI governance also highlights the need for a trust and trustworthiness perspective on AI systems, individuals, and institutions [2].

The rise of AI is impacting the use and management of digital trust. People's trust in AI is not just about the efficacy of the technology. It also relies on Transparency, Accountability, Human Oversight, Ethical Responsibility and Sound Governance Practices [2,4]. Yet, there is no existing comprehensive research on Responsible AI and AI governance. A systematic review indicates that there are multiple perspectives in the field, but a lack of integration in the governance areas [3]. Other recent research highlights the need for human oversight, documentation, data accountability, and the need to restrict algorithmic autonomy in the context of human-centred AI governance [4]. This means there is still a need for an understanding of how these governance areas can be integrated in a Human-Centric Approach with a focus on Digital Trust.

For this reason, this research carries out a systematic literature review on Responsible AI in a Data-Driven Society. The review aims to compile all the research work published or coordinated to highlight the most relevant governance areas, investigate their connection with Human-Focused Governance and Digital Trust, and showcase concepts, theories, and gaps in the literature when it comes to governance.

Responsible, Human-Centric Governance encompasses key elements of Privacy Protection, Transparency, Accountability, Ethical AI Governance, and Human Monitoring. The study seeks to shed light on the opportunities and challenges of how Responsible AI can contribute to the creation of trustworthy digital environments by bridging these fields.

Problem Statement

There is a wealth of information available on Data Governance, Responsible AI, Privacy, Transparency, Accountability and Digital Trust, but it is widely dispersed. There is only limited research that ties these together in one Human-Centric Data Governance approach to illustrate how responsible governance helps to achieve Digital Trust.

Research Methodology

Research Design

This study employs the Systematic Literature Review (SLR) approach as a research method to find, assess, and combine the results of previous research on the concept of Human-Centric Governance, Responsible AI, and Digital Trust in Data-Driven Societies. Systematic reviews are a methodical means of summarizing the existing evidence and highlighting gaps in knowledge that can inform future research.

The review process follows the PRISMA 2020 reporting guidelines [5] and comprises of five stages: Study Identification, Screening, Eligibility, Data Extraction, and Evidence Synthesis.

Questions: Research

RQ1: What are the research areas and governance aspects that are predominant for Human-Focused Governance and Digital Trust?

RQ2: What is the understanding of Privacy Protection, Transparency, Accountability, and Ethical AI Governance from the literature?

RQ3: What are the gaps in current approaches to governance in theory, method and practice?

RQ4: How to connect these governance dimensions conceptually to explain Digital Trust?

RQ5: Which relationships found in the literature can be restated in terms of testable hypotheses for empirical study?

Inclusion and Exclusion Criteria

Pre-defined inclusion and exclusion criteria will assure a consistent and transparent study selection. Reporting in systematic reviews is significant, and eligibility is important [5,6]. The Inclusion and Exclusion Criteria for this research are shown in Table 1.

Table 1. Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Peer-reviewed scholarly publications	Non-peer-reviewed/non-scholarly material
Research addressing Data Governance	Studies unrelated to Data Governance
Human-Centric / Human-Centered AI	Purely technical AI algorithm studies
Responsible / Ethical AI Governance	Studies with no governance dimension
Privacy, Transparency or Accountability	Studies unrelated to responsible data use
Digital Trust / Trust in AI	General technology adoption studies without trust relevance
AI-enabled Data-Driven Environments	Non-Digital/Non-AI contexts
Conceptual, empirical, or systematic research	Editorials, opinion pieces, and duplicate records
English-language publications	Non-English publications

Data Extraction and Synthesis

Eligible studies will systematically gather relevant information through a structured framework for data extraction. Systematic reviews should include characteristics of the studies included and the methods used to synthesize the results [5,6].

The following information will be extracted:

- Authors & Publication Year
- Research Focus
- Governance Dimension(s)
- Theoretical Perspective
- Research Methodology
- Key Findings
- Relationship with Digital Trust
- Identified Limitations and Research Gaps

This evidence will be categorized by theme to look for patterns, similarities, differences and gaps in the literature. A special emphasis will be given to the interactions between Privacy Protection, Transparency, Accountability, Ethical AI Governance, Human Monitoring and Digital Trust.

This synthesis will help to create a more complete understanding of Human-Centric Governance for Digital Trust in Responsible AI and Analytics-Driven Societies.

Literature Review — Chronological & Thematic Synthesis

The Chronological and Thematic Synthesis of Recent Literature on Human-Centric Governance, Responsible AI, and Digital Trust is summarized in Table 2.

Table 2. Literature Synthesis on Human-Centric Governance, Responsible AI, and Digital Trust

Year	Authors	Research Focus / Paper	Key Contribution	Identified Gap
2026	Lee, H.-K., Yoon, C., & Lee, B. G.	Operationalising Human-Centred AI Governance Under the EU AI Act [4]	Highlights Human Oversight, Data Accountability, Documentation, and Bounded Algorithmic Autonomy in Human-Centred AI Governance.	Limited Integration of Human-Centred Governance Mechanisms with Digital Trust.
2025	Papagiannidis, E., Mikalef, P., & Conboy, K.	Responsible Artificial Intelligence Governance: A Review and Research Framework [1]	Identifies Structural, Relational, and Procedural practices as important components of Responsible AI Governance.	Governance practices require stronger integration with Human-Centric Trust perspectives.
2025	Batool, A., Zowghi, D., & Bano, M.	AI Governance: A Systematic Literature Review [3]	Maps the Emerging AI-Governance literature and demonstrates its diverse research perspectives.	AI-governance research remains fragmented, with limited comprehensive integration.
2024	Lahusen, C., Maggetti, M., & Slavkovik, M.	Trust, Trustworthiness and AI Governance [2]	Examines trust and trustworthiness across AI Systems, Humans, and Institutions.	Trust and Governance Dimensions are not sufficiently integrated into a Unified Framework.
2024	Manzini, A., Keeling, G., Marchal, N., McKee, K. R., Rieser, V., & Gabriel, I.	Should Users Trust Advanced AI Assistants? Justified Trust as a Function of Competence and Alignment [7]	Examines justified trust in advanced AI Assistants through Competence and Alignment.	Trust requires broader consideration of Governance and Human-Centric dimensions.
2024	Novelli, C., Taddeo, M., & Floridi, L.	Accountability in Artificial Intelligence: What It Is and How It Works [8]	Establishes Accountability as a fundamental component of Responsible AI and examines how accountability operates in AI systems.	Accountability needs stronger integration with other Governance Dimensions and Digital Trust.

Critical Synthesis and Research Gaps

What the Literature Establishes

The literature indicates that Responsible AI Governance requires mechanisms for Privacy, Transparency, Accountability, Ethical Responsibility, and Human Monitoring. Recent studies also show that trust and trustworthiness are important social and technical aspects of AI governance [1–4].

What Remains Fragmented

Although well researched, these areas are usually examined individually or separately. There is a lack of understanding of the interconnections among Responsible AI Governance, Privacy, Transparency, Accountability, Human Monitoring and Digital Trust, and these are sometimes treated as separate entities.

Identified Research Gaps

The identified research gaps in the Human-Centric Governance and Digital Trust Literature are depicted in Table 3.

Table 3. Identified Research Gaps in Human-Centric Governance and Digital Trust Literature

Gap Type	Identified Gap
Conceptual Gap	Lack of an integrated conceptualization connecting Human-Centric Governance with Digital Trust.
Integration Gap	Privacy, Transparency, Accountability, Ethical AI Governance, and Human Oversight are often studied independently.
Theoretical Gap	Limited explanation of how multiple governance mechanisms collectively contribute to Digital Trust.
Methodological Gap	Limited empirical operationalization of integrated Human-Centric Governance constructs.
Practical Gap	Need for governance approaches that translate Responsible AI principles into actionable Human-Centred practices.

Research Opportunity

The gaps identified in the research indicate the need for an integrated Human-Centric Governance approach that links responsible governance to Digital Trust within Data-Driven Societies.

Human-Centric Governance Framework for Digital Trust

Conceptual Foundation

According to the literature, Privacy Protection, Transparency, Accountability, Ethical AI Governance and Human Monitoring are the key elements of Responsible and People-Centred AI Governance [1-4]. But

these areas are usually investigated separately. In this study, it is suggested to embed them into a Human-Centric Governance approach, where Digital Trust is the main goal [7].

Proposed Framework

Figure 1 shows the Proposed Framework, considering Trustworthy Digital Transformation as an output.



Figure 1. Core Research Logic for Human-Centric Governance and Digital Trust

Proposed Relationships

Table 4. Proposed Hypothesized Relationships Between Human-Centric Governance Dimensions and Digital Trust

Relationship	Conceptual Proposition
Privacy Protection → Digital Trust	Effective protection and responsible handling of data can strengthen confidence in AI-enabled systems.
Transparency → Digital Trust	Greater transparency can improve users' understanding and confidence in AI-enabled decision-making.
Accountability → Digital Trust	Clearly defined responsibility for AI decisions and outcomes can strengthen perceived reliability and legitimacy.
Ethical AI Governance → Digital Trust	Governance aligned with ethical principles can strengthen perceptions of responsible and trustworthy AI.
Human Oversight → Digital Trust	Meaningful human oversight can reinforce human agency, responsibility, and confidence in AI systems.
Digital Trust → Trustworthy Digital Transformation	Greater confidence in AI-enabled systems can support their responsible and sustainable adoption.

Conceptual Contribution

The proposed framework shifts the focus from separate Responsible AI principles to an Integrated Human-Centric Governance perspective, in which multiple governance areas collaborate to build Digital Trust.

Discussion and Implications

Discussion

The systematic review shows that Responsible AI Governance now extends beyond technical issues to encompass Human Values, Accountability, Transparency, Privacy, and Trust [1–4]. Recent research also shows that building the right kind of trust in AI requires considering Transparency, Explainability, Reliability, Privacy, Human Autonomy, and Agency, rather than assuming that trust is inherently good [8–9].

The literature also indicates that trustworthy AI-enabled decision-making requires multiple governance dimensions rather than dependence on a single technical or regulatory mechanism. Agbabiaka et al. [10], for example, identify multiple trustworthiness requirements and accentuate the significance of accountability in AI-enabled automated decision-making.

Similarly, AI governance research emphasizes embedding Trustworthiness, Transparency, Human Monitoring, and Accountability into the development and deployment of AI systems [10].

The proposed framework is connected in these areas with a Human-Centric Governance perspective with Digital Trust at the centre. This is in line with the recent research on the Human-centred AI, which has been centered on the integration of AI systems into human values, needs and social aspects [11].

Digital Trust goes beyond technical reliability, as the synthesis suggests. Instead, it should be thought of as a governance outcome and as a consequence of its Privacy, Transparency, Accountability, Ethical Responsibility, and Meaningful Human Engagement [12].

Conceptual Consequences

1. Including the divided dimensions of governance in a Human-centric Governance view [1,3,11].
2. Proposing a framework for the establishment of Digital Trust [3-5].
3. Associate Responsible AI Governance with ideas of Human-Centred Governance principles [1,10,11].

4. Exploring the link between Privacy, Transparency, Accountability, Ethical AI Governance, Human Monitoring and Digital Trust and a Conceptual Foundation to study the relationship between these [4,8,9].

Applied Consequences

- Policymakers in developing Human-Centred AI Governance Policies [9,10].
- Organizations in the design of Accountable & Transparent AI Governance mechanisms [1,10].
- Lack of integration of Privacy, Explainability, Accountability, and Human Monitoring in AI systems by the developers [4,10].
- Institutions should strengthen appropriate Digital Trust rather than maximize users' trust in AI [8].

Consequences for Data-Driven Societies

The study reveals that Responsible Digital Transformation should involve Governance Approaches that prioritize People, Responsibility, Transparency, and Trust in the implementation of AI. The Human-Centred AI research also highlights the need to ensure AI development is in line with human needs and social values [11].

Conclusions and Future Research Directions

Conclusions

- This study was conducted to look at the fragmentation of research on Responsible AI Governance, Human-Centric Governance and Digital Trust, with a specific interest in the comparatively small number of studies that combined the five aspects of Privacy Protection, Transparency, Accountability, Ethical AI Governance and Human Monitoring.
- The systematic approach used was a SLR with steps as described in the PRISMA 2020 systematic process [5-6] and relevant to the literature.
- Key Findings: Privacy Protection, Transparency, Accountability, Ethical AI Governance, and Human Monitoring are identified as key components of Responsible AI Governance. But, these regions are still poorly studied today.
- Limitations and Future Work: The proposed relationships are yet to be validated with actual data.

Key Contribution

1. The study brings together diverse literature on Responsible AI, Human-Centric Governance and Digital Trust and creates a research perspective for Data-Driven Societies.
2. Based on the review, Thematic Synthesis of Governance Dimensions, the review characteristically identifies and synthesizes Privacy Protection, Transparency, Accountability, Ethical AI Governance and Human Monitoring as interlinked dimensions of responsible Human-Centric Governance.

3. Identification of the Governance – Trust Relationship: The study highlights the importance of Digital Trust as an outcome of responsible governance and not just a technological or user-perception problem.
4. The research gap consolidation showed that there is a gap between theory and practice in the conceptual, theoretical, methodological and integration parts of the previous research.
5. Upon reviewing the Literature, this study proposes a conceptual research design, shown in Figure 2.

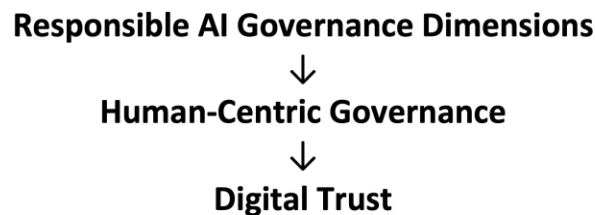


Figure 2. Conceptual Research Direction: Integrated Pathway

References

1. E. Papagiannidis, P. Mikalef, and K. Conboy, "Responsible artificial intelligence governance: A review and research framework," *Journal of Strategic Information Systems*, vol. 34, no. 2, Art. no. 101885, 2025.
<https://doi.org/10.1016/j.jsis.2024.101885>
2. C. Lahusen, M. Maggetti, and M. Slavkovik, "Trust, trustworthiness and AI governance," *Scientific Reports*, vol. 14, Art. no. 20752, 2024.
<https://doi.org/10.1038/s41598-024-71761-0>
3. A. Batool, D. Zowghi, and M. Bano, "AI governance: A systematic literature review," *AI and Ethics*, vol. 5, pp. 3265–3279, 2025.
<https://doi.org/10.1007/s43681-024-00653-w>
4. H.-K. Lee, C. Yoon, and B. G. Lee, "Operationalising human-centred AI governance under the EU AI Act: A governance framework for human oversight and data accountability," *Systems*, vol. 14, no. 7, Art. no. 849, 2026.
<https://doi.org/10.3390/systems14070849>
5. M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, Art. no. n71, 2021.
<https://doi.org/10.1136/bmj.n71>
6. M. J. Page *et al.*, "PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews," *BMJ*, vol. 372, Art. no. n160, 2021.
<https://doi.org/10.1136/bmj.n160>
7. A. Manzini, G. Keeling, N. Marchal, K. R. McKee, V. Rieser, and I. Gabriel, "Should users trust advanced AI assistants? Justified trust as a function of competence and alignment," in *Proc. 2024 ACM Conf. Fairness, Accountability, Transparency (FAccT)*, 2024, pp. 1174–1186.
<https://doi.org/10.1145/3630106.3658964>
8. C. Novelli, M. Taddeo, and L. Floridi, "Accountability in artificial intelligence: What it is and how it works," *AI and Society*, vol. 39, pp. 1871–1882, 2024.
<https://doi.org/10.1007/s00146-023-01635-y>

9. S. Mehrotra, C. Degachi, O. Vereschak, C. M. Jonker, and M. L. Tielman, "A systematic review on fostering appropriate trust in human-AI interaction: Trends, opportunities and challenges," *ACM Journal on Responsible Computing*, vol. 1, no. 4, Art. no. 26, 2024.
<https://doi.org/10.1145/3696449>
10. O. Agbabiaka, A. Ojo, and N. Connolly, "Requirements for trustworthy AI-enabled automated decision-making in the public sector: A systematic review," *Technological Forecasting and Social Change*, vol. 215, Art. no. 124076, 2025.
<https://doi.org/10.1016/j.techfore.2025.124076>
11. S. P. Poti and C. J. Stanton, "Enabling affordances for AI governance," *Journal of Responsible Technology*, vol. 18, Art. no. 100086, 2024.
<https://doi.org/10.1016/j.jrt.2024.100086>
12. "Understanding human-centred AI: A review of its defining elements and a research agenda," *Behaviour & Information Technology*, vol. 44, pp. 3771–3810, 2025.
<https://doi.org/10.1080/0144929X.2024.2448719>