

Humanizing Data Science Through Domain Expertise for Ethical and Trustworthy Artificial Intelligence

Marviola Hardini¹ , Nur Azizah² , Syahla Naurah³ , Rendhika Adyatama^{4*} ,
Souza Nurafrianto Windiartono Putra⁵

¹Indonesian Lecturers Association, Indonesia

^{2,3,4}Faculty of Economics and Business, University of Raharja, Indonesia

⁵Department of Information Technology, Universitat des Saarlandes, Germany

¹marviola@raharja.info, ²nur.azizah@raharja.info, ³syahla@raharja.info, ⁴rendhika@raharja.info, ⁵s.windiartono@uni-saarland.de

*Corresponding Author

Article Info

Article history:

Submission June 10, 2026
Revised July 7, 2026
Accepted July 23, 2026
Published August 12, 2026

Keywords:

Humanizing Data Science
Domain Expertise
Ethical Artificial Intelligence
Trustworthy AI
Systematic Literature Review



ABSTRACT

The rapid adoption of Artificial Intelligence (AI) has transformed data-driven decision making across healthcare, finance, education, and public services. Despite these advances, AI systems continue to face challenges related to algorithmic bias, limited contextual understanding, insufficient transparency, and declining user trust, highlighting the need to integrate human expertise throughout the data science process. **This study aims** to systematically examine the role of domain expertise in Humanizing Data Science for the development of Ethical and Trustworthy Artificial Intelligence. **A Systematic Literature Review (SLR)** was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Peer-reviewed studies published between 2021 & 2026 were identified, screened using predefined inclusion and exclusion criteria, and analyzed through thematic synthesis. **The review identifies** five major themes, namely Human-Centered Artificial Intelligence, Domain Expertise Integration, Ethical and Trustworthy AI, Explainability and Human Oversight, and Humanizing Data Science. Based on these findings, this study proposes a Humanizing Data Science Framework integrating domain expertise throughout the AI lifecycle. **The framework** demonstrates that combining technical capabilities with human knowledge supports transparent, fair, accountable, trustworthy, and human-centered AI, providing valuable practical and theoretical guidance for researchers, practitioners, organizations, policymakers, and future interdisciplinary innovation initiatives worldwide.

This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.



This is an open-access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>)
©Authors retain all copyrights

1. INTRODUCTION

Artificial Intelligence (AI) has become a key driver of digital transformation, enabling organizations to automate processes, generate predictive insights, and support complex decision making across healthcare, finance, education, manufacturing, transportation, and public services. Beyond improving operational efficiency and innovation, AI also contributes to the United Nations Sustainable Development Goals (SDGs), particularly SDGs 3 (Good Health and Well-being), SDGs 9 (Industry, Innovation and Infrastructure), and SDG 16 (Peace, Justice and Strong Institutions) [1]. As AI continues to shape modern society, expectations have shifted from developing highly accurate algorithms toward designing intelligent systems that respect human values, promote fairness, transparency, accountability, inclusiveness, and human well-being. This shift has encouraged

the emergence of Human Centered Artificial Intelligence (HCAI), which emphasizes that AI should augment rather than replace human capabilities [2, 3].

Despite these advances, the rapid adoption of AI has introduced significant ethical and societal challenges, including biased predictions, opaque decision-making, and context-insensitive recommendations that undermine public trust and responsible AI adoption. These issues often result from AI models relying primarily on historical data and computational optimization while overlooking contextual knowledge, professional expertise, and human judgment that cannot be fully captured by data alone. In high-impact domains such as healthcare, finance, law, and public administration, these limitations directly affect human welfare, equity, and social justice. Consequently, technological sophistication alone is insufficient to ensure trustworthy AI, highlighting the need for Humanizing Data Science as an approach that integrates domain expertise throughout the data science lifecycle to improve contextual understanding, reduce unintended bias, support ethical reasoning, and produce socially beneficial decisions [4].

Recent research on Ethical AI, Explainable AI, Responsible AI, and Trustworthy AI has substantially improved understanding of transparency, fairness, interpretability, governance, and regulatory compliance. However, relatively limited attention has been given to the role of domain expertise in humanizing data science. In many AI projects, domain experts participate mainly during model validation or implementation, while their involvement in problem formulation, data preparation, feature engineering, model interpretation, and continuous evaluation remains insufficiently explored [5, 6]. As a result, knowledge regarding the integration of domain expertise across the AI lifecycle remains fragmented, creating an important research gap for organizations seeking AI systems that combine technical excellence with meaningful societal impact.

To address this gap, this study conducts a SLR following the PRISMA guidelines to synthesize current evidence on the role of domain expertise in Humanizing Data Science for ethical and trustworthy AI. The review identifies the major roles of domain expertise, examines challenges in integrating human expertise into AI development, and analyzes its influence on transparency, accountability, contextual intelligence, stakeholder trust, and human well-being [7].

Furthermore, this study proposes a Humanizing Data Science Framework that illustrates how domain expertise can be integrated throughout the AI lifecycle to strengthen ethical decision making and support responsible innovation. The proposed framework contributes theoretically by extending Human Centered Artificial Intelligence through the integration of human knowledge and data science, while providing practical guidance for researchers, AI developers, industry practitioners, and policymakers. Ultimately, this study supports the realization of SDGs 3, SDGs 9, and SDGs 16 by promoting AI systems that are not only technologically advanced but also ethical, trustworthy, inclusive, and centered on human well-being [8].

1.1. Human Centered Artificial Intelligence

AI has become a key technology supporting data-driven automation and intelligent decision making across diverse sectors. While early AI research emphasized computational efficiency and predictive performance, recent developments increasingly focus on aligning AI with human values and ethical principles. This shift has led to HCAI, which places humans at the center of AI design, development, deployment, and evaluation. HCAI promotes collaboration between humans and AI by combining computational capabilities with contextual understanding, ethical reasoning, and professional judgment [9]. Consequently, AI systems are expected to achieve not only high performance but also fairness, explainability, accountability, robustness, and trustworthiness.

Beyond technical performance, HCAI seeks to ensure that AI contributes to sustainable development and human well-being, supporting SDGs 3 (Good Health and Well-being), SDGs 9 (Industry, Innovation and Infrastructure), and SDGs 16 (Peace, Justice and Strong Institutions) through responsible, transparent, and inclusive AI. Accordingly, AI is increasingly evaluated based on both its technical capabilities and societal impact. Nevertheless, existing HCAI research primarily emphasizes human-centered design, governance, and human-computer interaction, while providing limited attention to integrating domain expertise throughout the data science lifecycle. Consequently, Humanizing Data Science extends HCAI by embedding domain expertise across AI development to produce context-aware, ethical, and trustworthy AI systems [10].

1.2. Humanizing Data Science

The rapid advancement of data science has transformed organizational decision making through machine learning, predictive analytics, and generative AI. Traditionally, data science has emphasized predictive accuracy, computational efficiency, and model performance. However, real-world AI applications reveal that

data alone cannot fully represent contextual knowledge, professional experience, or tacit expertise. Because datasets may be incomplete, context-dependent, or influenced by historical bias, AI systems optimized solely through data-driven approaches may achieve high performance while failing to address practical realities, ethical considerations, and human-centered needs, particularly in critical domains such as healthcare, finance, education, and public administration [11, 12]. These challenges have increased interest in Humanizing Data Science, which advocates integrating human knowledge throughout the AI lifecycle and builds upon related concepts in Human-Centered Artificial Intelligence, Human-in-the-Loop systems, collaborative intelligence, Responsible AI, and socio-technical systems that consistently emphasize meaningful human participation [13].

Accordingly, Humanizing Data Science embeds domain expertise, contextual understanding, ethical reasoning, and interdisciplinary collaboration into problem formulation, data preparation, model development, validation, deployment, and continuous evaluation, thereby improving contextual awareness, reducing bias, and strengthening the relevance, interpretability, and trustworthiness of AI outcomes [14]. Despite growing recognition of these principles, existing research remains fragmented, with human oversight, explainability, governance, and ethical AI often examined separately, while the role of domain expertise as the primary mechanism for Humanizing Data Science has not been systematically synthesized. Consequently, Humanizing Data Science should be viewed as an interdisciplinary paradigm that integrates technical capabilities with human expertise to support the development of ethical, trustworthy, and socially responsible AI, providing the conceptual foundation for the subsequent discussion on domain expertise in AI development [15].

1.3. Domain Expertise in AI Development

Domain expertise refers to the specialized knowledge, professional experience, and contextual understanding that enable decisions beyond what can be derived from data alone. In AI development, domain experts contribute tacit knowledge, regulatory insights, ethical considerations, and contextual understanding throughout the AI lifecycle. Their roles include defining practical objectives, supporting data preparation, guiding model development and interpretation, reducing bias, and participating in evaluation and post-deployment monitoring to ensure AI systems remain adaptive, ethical, and aligned with organizational needs [15, 16]. Recent studies further demonstrate that integrating domain expertise enhances AI performance, reliability, contextual intelligence, explainability, and societal value while strengthening interdisciplinary collaboration, stakeholder trust, and AI-assisted decision making, particularly in high-stakes domains such as healthcare, finance, education, legal services, and public administration [17].

Despite its recognized importance, the integration of domain expertise remains inconsistent across AI research and practice, as many frameworks continue to prioritize algorithmic optimization while providing limited guidance for involving experts throughout the data science lifecycle. Consequently, domain expertise should be recognized as a strategic and fundamental pillar of Humanizing Data Science, providing the contextual foundation for fairness, transparency, accountability, and the development of Ethical and Trustworthy Artificial Intelligence rather than serving merely as a resource introduced after model development [18].

1.4. Ethical and Trustworthy Artificial Intelligence

The widespread adoption of AI has intensified concerns regarding the ethical and societal implications of AI-driven decision making. Although advances in machine learning have improved predictive capabilities, challenges such as algorithmic bias, limited transparency, privacy risks, and accountability continue to hinder responsible AI deployment [19]. These concerns have led to the emergence of Ethical AI and Trustworthy AI, which promote fair, transparent, explainable, accountable, and reliable AI systems aligned with human values. International frameworks, including the OECD AI Principles, the European Commission's Ethics Guidelines for Trustworthy AI, the NIST AI Risk Management Framework, and ISO/IEC 42001, emphasize the importance of effective governance and meaningful human involvement [20, 21]. Ethical AI embeds ethical principles throughout the AI lifecycle, whereas Trustworthy AI ensures that AI systems remain reliable, safe, explainable, and compliant with legal requirements.

Despite these advances, many organizations continue to prioritize predictive performance over the early integration of ethical considerations, creating a gap between ethical principles and practical implementation. Domain expertise helps bridge this gap by providing contextual knowledge, professional judgment, and interdisciplinary collaboration throughout the data science lifecycle. Its integration improves risk identification, explainability, transparency, accountability, and stakeholder trust, reinforcing Humanizing Data Science as the foundation for developing ethical, trustworthy, and human-centered AI [22].

1.5. Conceptual Synthesis and Research Gap

Previous studies have highlighted HCAI, Domain Expertise, and Ethical and Trustworthy Artificial Intelligence as key foundations of responsible AI development. Although existing research has advanced human-centered design, AI governance, transparency, fairness, explainability, accountability, and the role of human expertise in improving AI performance and decision quality while reducing bias, these topics are generally examined independently [23]. Consequently, the systematic integration of domain expertise throughout the data science lifecycle remains fragmented. Table 1 summarizes the major research themes, existing limitations, and the contribution proposed by this study.

Table 1. Research Gap Analysis

Research Area	Existing Literature	Identified Gap	Contribution of This Study
Human Centered Artificial Intelligence	Focuses on human-centered design, usability, and collaboration between humans and AI.	Limited discussion regarding the operational implementation of human-centered principles throughout the data science lifecycle.	Explains how Human Centered AI can be operationalized through the Humanizing Data Science approach.
Humanizing Data Science	Related concepts are discussed through Human-in-the-Loop, socio-technical systems, and collaborative intelligence.	No integrated conceptual framework specifically explaining Humanizing Data Science.	Synthesizes multidisciplinary evidence to establish Humanizing Data Science as a comprehensive conceptual perspective.
Domain Expertise	Domain experts are recognized as essential within specific application areas.	Existing studies rarely describe a systematic integration of domain expertise across all stages of AI development.	Identifies the strategic role of domain expertise throughout the entire AI lifecycle.
Ethical and Trustworthy AI	Emphasizes fairness, explainability, accountability, transparency, and governance.	Existing research pays limited attention to how domain expertise contributes to implementing these principles in practice.	Demonstrates that domain expertise serves as a practical mechanism for achieving Ethical and Trustworthy AI.
Integrated Perspective	Previous studies investigate HCAI, Domain Expertise, and Ethical AI as separate research topics.	A comprehensive framework integrating these concepts remains unavailable.	Proposes the Humanizing Data Science Framework to integrate Human Centered AI, Domain Expertise, and Ethical & Trustworthy AI.

As presented in Table 1, previous studies have made substantial contributions to Human Centered AI, Domain Expertise, and Ethical and Trustworthy AI. Nevertheless, these studies remain fragmented and predominantly focus on individual aspects rather than providing an integrated perspective that connects these concepts within the data science lifecycle. This fragmentation indicates a theoretical and practical gap in understanding how domain expertise can systematically humanize data science to support the development of ethical and trustworthy AI systems.

Accordingly, this study addresses the identified gap by conducting a SLR to synthesize multidisciplinary evidence and develop a Humanizing Data Science Framework. The proposed framework positions domain expertise as the central mechanism that links Human Centered AI principles with Ethical and Trustworthy Artificial Intelligence, thereby providing a holistic perspective for designing AI systems that are technically reliable, ethically responsible, and capable of generating positive societal impacts in support of SDGs 3, SDGs 9, and SDGs 16.

2. RESEARCH METHODOLOGY

2.1. Research Design

This study employed a SLR to synthesize existing knowledge on the role of domain expertise in Humanizing Data Science for the development of Ethical and Trustworthy Artificial Intelligence. An SLR was selected because it provides a structured, transparent, and reproducible approach for identifying, evaluating, and synthesizing scientific evidence from multiple disciplines while minimizing selection bias through predefined review procedures [24, 25]. The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines, which ensure methodological rigor through four sequential stages: identification, screening, eligibility assessment, and study inclusion.

The review focused on peer-reviewed publications addressing Human Centered Artificial Intelligence, Humanizing Data Science, Domain Expertise, Ethical AI, Trustworthy AI, Human-in-the-Loop systems, Explainable AI, and Responsible AI. The selected studies were analyzed to identify recurring themes, conceptual relationships, and evidence regarding the integration of domain expertise throughout the AI development life-cycle [26].

2.2. Review Protocol

The review protocol followed the PRISMA 2020 framework to ensure a systematic and reproducible study selection process consisting of four sequential stages, namely identification, screening, eligibility assessment, and inclusion. During the identification stage, relevant publications were retrieved from selected scientific databases using predefined search keywords combined with Boolean operators. The screening stage involved removing duplicate records and evaluating article titles and abstracts for relevance, followed by full-text assessment against predefined inclusion and exclusion criteria during the eligibility stage [27]. Only studies meeting all requirements were included in the final qualitative synthesis and thematic analysis. This framework provides a structured and transparent procedure that enhances the methodological rigor, reliability, and reproducibility of the review findings.

2.3. Search Strategy

A comprehensive search strategy was developed to identify studies on the integration of domain expertise in Humanizing Data Science for Ethical and Trustworthy Artificial Intelligence. The literature search used major scientific databases indexing peer-reviewed publications in computer science, artificial intelligence, information systems, and interdisciplinary research. The search included English-language journal articles and conference proceedings published between 2021 & 2026 [28]. Keywords were derived from the core concepts of this study and combined using the Boolean operators AND and OR to improve search precision. The complete search strategy and selected databases are presented in Table 2.

Table 2. Search Databases and Search Keywords

Database	Coverage
Scopus	Multidisciplinary peer-reviewed journals and conference proceedings
Web of Science	High-quality multidisciplinary scientific publications
IEEE Xplore	Engineering, Artificial Intelligence, and Computer Science
ACM Digital Library	Computing, Human-Computer Interaction, and AI research
ScienceDirect	Computer Science, Information Systems, and interdisciplinary studies
SpringerLink	Artificial Intelligence, Data Science, and Technology research

Table 2 presents the scientific databases selected for this review and their respective research coverage. Six internationally recognized databases were included to ensure broad disciplinary representation and comprehensive retrieval of high-quality publications. Scopus and Web of Science provide extensive multidisciplinary coverage of peer-reviewed journals and conference proceedings, while IEEE Xplore and ACM Digital Library capture research in computer science, artificial intelligence, software engineering, and human-computer interaction. ScienceDirect and SpringerLink further broaden the search to interdisciplinary studies in data science, information systems, and emerging AI technologies. Collectively, these databases maximize literature coverage, reduce publication bias, and strengthen the reliability of the evidence synthesized in this review.

Based on the scope of these databases, a comprehensive search strategy was developed using keywords representing the principal concepts of this study, namely Humanizing Data Science, Human-Centered

Artificial Intelligence, Domain Expertise, Ethical AI, Trustworthy AI, Human-in-the-Loop, Explainable AI, and Responsible AI [29]. These keywords were combined using the Boolean operators AND and OR to improve search precision while ensuring comprehensive retrieval of relevant publications. The Boolean search string used in this review is presented below.

```
("Humanizing Data Science" OR "Human Centered AI" OR "Human Centered Artificial Intelligence")
AND
("Domain Expertise" OR "Domain Knowledge" OR "Expert Knowledge")
AND
("Ethical AI" OR "Trustworthy AI" OR "Responsible AI" OR "Explainable AI" OR "Human in the Loop")
```

The selected databases and search strategy were designed to maximize the retrieval of relevant publications while minimizing irrelevant records. Furthermore, using multiple databases reduces publication bias and enhances the comprehensiveness of the systematic review. The retrieved records were subsequently processed through the PRISMA 2020 workflow, including duplicate removal, title and abstract screening, full-text eligibility assessment, and final study inclusion for qualitative synthesis [30].

2.4. Inclusion and Exclusion Criteria

To ensure the quality, relevance, and consistency of the reviewed studies, predefined inclusion and exclusion criteria were established before the screening process. These criteria were designed to select publications that directly addressed the objectives of this review while excluding studies that did not provide sufficient empirical or conceptual evidence related to Humanizing Data Science, Domain Expertise, or Ethical and Trustworthy Artificial Intelligence [31, 32].

Table 3. Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion
Publication Year	2021–2026	Before 2021
Language	English	Non-English publications
Publication Type	Peer-reviewed journal articles and conference proceedings	Books, theses, dissertations, editorials, reports, and unpublished manuscripts
Research Focus	Human Centered AI, Humanizing Data Science, Domain Expertise, Ethical AI, Trustworthy AI, Explainable AI, Human-in-the-Loop, and Responsible AI	Studies unrelated to AI or lacking relevance to the research objectives
Accessibility	Full-text available	Abstract only or inaccessible full text

Table 3 summarizes the inclusion and exclusion criteria applied during the study selection process. These criteria were established to ensure that only high-quality and relevant publications addressing the objectives of this review were included. The selection focused on studies published between 2021 & 2026 concerning Human-Centered Artificial Intelligence, Humanizing Data Science, Domain Expertise, Ethical AI, and Trustworthy AI. Only English-language peer-reviewed journal articles and conference proceedings with accessible full-text versions were included, whereas books, theses, editorials, reports, unpublished manuscripts, and studies unrelated to the research objectives were excluded.

The predefined criteria ensured a consistent, transparent, and reproducible screening process while minimizing the inclusion of irrelevant or low-quality studies. Consequently, the selected literature provides a reliable foundation for the subsequent thematic synthesis and the development of the proposed Humanizing Data Science Framework [33].

2.5. Study Selection Process

The study selection process followed the PRISMA 2020 framework to ensure a transparent and systematic identification of relevant literature. Publications retrieved from the six selected scientific databases using the predefined search strategy were merged into a single database for duplicate removal, followed by title

and abstract screening to exclude studies that were not relevant to the review objectives [1]. The remaining publications underwent full-text eligibility assessment based on the predefined inclusion and exclusion criteria, and only studies meeting all requirements were included in the final qualitative synthesis. The complete study selection process is illustrated in Figure 1.

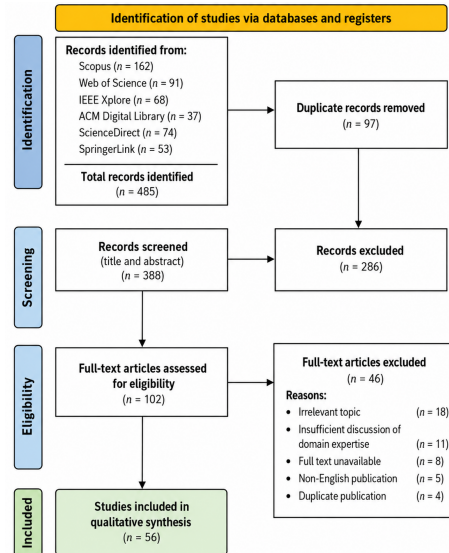


Figure 1. PRISMA Flow Diagram for Study Selection

As illustrated in Figure 1, the initial search across six scientific databases identified a total of 485 records. After removing 97 duplicate publications, 388 records remained for title and abstract screening. A total of 286 studies were excluded because they did not meet the research objectives. Subsequently, 102 full-text articles were assessed for eligibility, resulting in the exclusion of 46 studies based on the predefined inclusion and exclusion criteria. Finally, 56 studies were included in the qualitative synthesis and served as the primary evidence for developing the Humanizing Data Science Framework for Ethical and Trustworthy Artificial Intelligence.

2.6. Data Extraction and Analysis

Following the study selection process, the included publications were systematically analyzed through a structured data extraction procedure. A standardized extraction form was developed to ensure consistency in collecting relevant information from each study. The extracted information included bibliographic details, research objectives, application domains, research methods, AI approaches, the role of domain expertise, ethical and trustworthy AI dimensions, key findings, and identified research gaps.

Table 4. Data Extraction Items

Data Category	Description
Bibliographic Information	Author(s), publication year, journal or conference
Research Objective	Main purpose of the study
Research Domain	Healthcare, finance, education, manufacturing, public services, etc.
Research Method	Empirical study, case study, conceptual paper, review, or framework
AI Approach	Machine Learning, Deep Learning, Generative AI, Explainable AI, Human-in-the-Loop, etc.
Domain Expertise	Type and role of expert involvement in AI development
Ethical and Trustworthy AI Aspects	Fairness, transparency, accountability, explainability, privacy, governance, or trust

Key Findings	Main results reported by the study
Research Gap	Limitations and future research directions identified by the authors

Table 4 presents the results of the PRISMA-based study selection process. The initial search identified 325 records, of which 42 duplicate publications were removed, leaving 283 records for title and abstract screening. Following this stage, 89 articles were eligible for full-text assessment, and 56 studies satisfied all inclusion criteria and were retained for qualitative synthesis and thematic analysis.

The progressive reduction in the number of studies reflects the rigorous application of the predefined selection criteria, ensuring that only high-quality and relevant publications on Human-Centered Artificial Intelligence, Humanizing Data Science, Domain Expertise, and Ethical and Trustworthy AI were included. This systematic process enhances the transparency, reliability, and reproducibility of the review while minimizing selection bias.

3. RESULTS AND DISCUSSION

3.1. Characteristics of Selected Studies

Following the PRISMA 2020 screening procedure, 56 studies were selected for qualitative synthesis. Published between 2021 & 2026, these studies cover HCAI, Humanizing Data Science, Domain Expertise, Explainable AI, Responsible AI, and Ethical and Trustworthy Artificial Intelligence. The selected literature includes empirical studies, conceptual papers, systematic literature reviews, and framework development studies, providing multidisciplinary perspectives on integrating human expertise into AI while addressing ethical, technical, and organizational challenges. Table 5 summarizes the characteristics of the selected studies by publication year and research type.

Table 5. Characteristics of Selected Studies ($n = 56$)

Characteristic	Category	Frequency	Percentage (%)
Publication Year	2021	6	10.7
	2022	8	14.3
	2023	12	21.4
	2024	15	26.8
	2025	11	19.6
	2026	4	7.2
Research Type	Empirical Study	27	48.2
	Systematic Literature Review	11	19.6
	Conceptual Paper	10	17.9
	Framework/Model Development	8	14.3
Total		56	100.0

As presented in Table 5, publications on Human-Centered Artificial Intelligence, Domain Expertise, and Ethical AI have increased substantially over the past five years, with more than two-thirds of the selected studies published between 2023 & 2025. This trend reflects the growing academic focus on transparent, explainable, ethical, and human-centered AI, while nearly half of the selected studies were empirical, indicating an increasing emphasis on validating AI applications in real-world contexts. The reviewed studies were further categorized by their primary application domains, as presented in Table 6.

Table 6. Distribution of Selected Studies by Application Domain

Application Domain	Frequency	Percentage (%)
Healthcare	15	26.8
Finance	9	16.1
Education	8	14.3
Manufacturing	7	12.5

Public Administration	6	10.7
Smart Cities	5	8.9
General AI Framework	6	10.7
Total	56	100.0

Table 6 shows that healthcare is the dominant application domain, accounting for 26.8% of the selected studies. This finding underscores the importance of domain expertise in AI-assisted clinical decision making, patient monitoring, diagnosis, and healthcare delivery. Finance and education also represent major application areas, reflecting growing attention to fairness, transparency, and accountability in AI-supported decision making. Studies classified under General AI Framework mainly address conceptual models, governance, and interdisciplinary approaches for Ethical and Trustworthy AI. Overall, the diversity of application domains demonstrates that integrating domain expertise is a cross-sectoral requirement, reinforcing Humanizing Data Science as a multidisciplinary paradigm.

3.2. Proposed Humanizing Data Science Framework

The thematic synthesis identified five interconnected themes explaining how Artificial Intelligence can be developed in a more human-centered, ethical, and trustworthy manner. Together, these themes form an integrated process in which domain expertise connects technological capabilities with human values. Based on the synthesis of the 56 selected studies, this research proposes the Humanizing Data Science Framework, which integrates contextual knowledge, ethical reasoning, human oversight, and collaborative intelligence throughout the AI lifecycle. The framework demonstrates that responsible AI depends not only on advanced algorithms but also on the continuous integration of human expertise to achieve Ethical and Trustworthy Artificial Intelligence.

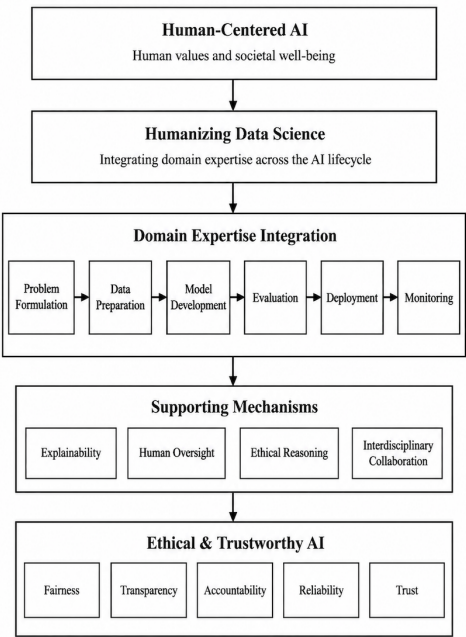


Figure 2. Humanizing Data Science Framework

Figure 2 illustrates that Human-Centered Artificial Intelligence provides the foundation for AI development by promoting technologies that enhance human capabilities and societal well-being. Within this framework, domain expertise integrates contextual knowledge, professional judgment, and ethical considerations throughout the AI lifecycle, while explainability and human oversight ensure transparent, interpretable, and accountable AI decisions. Together, these elements support the development of Ethical and Trustworthy Artificial Intelligence characterized by fairness, reliability, and stakeholder trust. The proposed framework unifies previously fragmented research by explaining how domain expertise operationalizes Human-Centered AI through Humanizing Data Science, providing practical guidance for AI developers, organizations, and policymakers to balance technological innovation with ethical responsibility and human well-being. Furthermore,

the framework supports SDGs 3 (Good Health and Well-being), SDGs 9 (Industry, Innovation and Infrastructure), and SDGs 16 (Peace, Justice and Strong Institutions) by promoting responsible, human-centered AI that generates sustainable societal value.

4. MANAGERIAL IMPLICATIONS

The findings of this study provide practical implications for organizations and AI practitioners involved in developing Artificial Intelligence systems. The proposed Humanizing Data Science Framework emphasizes the integration of domain expertise throughout the AI lifecycle, from problem formulation to deployment and continuous monitoring. This integration improves data quality, reduces algorithmic bias, enhances model transparency, and strengthens stakeholder trust while fostering interdisciplinary collaboration among data scientists, domain experts, and end users to develop AI systems that are technically accurate, contextually relevant, and explainable.

From a humanistic perspective, the proposed framework highlights that successful AI development should prioritize human values alongside technological advancement. Integrating ethical reasoning, human oversight, and domain expertise into data science practices supports the development of Ethical and Trustworthy Artificial Intelligence that promotes fairness, accountability, inclusivity, transparency, and societal well-being.

The findings also provide guidance for policymakers seeking to establish responsible AI governance that balances technological innovation with ethical responsibility. These implications align with the mission of the Journal of Orange Technology and contribute to the achievement of SDGs, particularly SDGs 3 (Good Health and Well-being), SDGs 9 (Industry, Innovation and Infrastructure), and SDGs 16 (Peace, Justice and Strong Institutions), by encouraging the adoption of responsible, trustworthy, and human-centered AI systems.

5. CONCLUSION

This study employed a SLR to examine the role of domain expertise in Humanizing Data Science for developing Ethical and Trustworthy Artificial Intelligence. Based on the synthesis of 56 peer-reviewed studies published between 2021 and 2026, five dominant themes were identified: Human-Centered Artificial Intelligence, Domain Expertise Integration, Ethical and Trustworthy AI, Explainability and Human Oversight, and Humanizing Data Science. The findings demonstrate that responsible AI development requires the continuous integration of human knowledge, contextual understanding, and ethical principles throughout the AI lifecycle.

This study addresses the fragmented understanding of domain expertise integration by proposing the Humanizing Data Science Framework, which positions domain expertise as the central mechanism linking human-centered design, explainability, ethical reasoning, and trustworthy AI. The framework contributes a unified conceptual model and practical guidance for researchers, AI practitioners, organizations, and policy-makers to develop AI systems that balance technological innovation with responsible governance and human well-being.

This study is limited by its reliance on a Systematic Literature Review, and the proposed framework has not yet been empirically validated in real-world settings. Future research should validate the framework using quantitative, qualitative, or mixed-method approaches across domains such as healthcare, education, finance, and public services. Such studies would strengthen the evidence for integrating domain expertise into Humanizing Data Science and further advance Ethical and Trustworthy Artificial Intelligence in supporting sustainable and human-centered digital transformation.

6. DECLARATIONS

6.1. About Authors

Marviola Hardini (MH)  <https://orcid.org/0000-0003-3336-2131>

Nur Azizah (NA)  <https://orcid.org/0000-0002-5333-5520>

Syahla Naurah (SN)  <https://orcid.org/0009-0003-2577-1995>

Rendhika Adyatama (RA)  <https://orcid.org/0009-0008-7071-5374>

Souza Nurafrianto Windiarsono Putra (SN)  -

6.2. Author Contributions

Conceptualization: SN; Methodology: MH; Software: SN; Validation: NA and RA; Formal Analysis: MH and RA; Investigation: SN; Resources: SN; Data Curation: MH; Writing Original Draft Preparation: RA and SN; Writing Review and Editing: NA and MH; Visualization: SN; All authors, MH, NA, SN, RA, and SN, have read and agreed to the published version of the manuscript.

6.3. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.4. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.5. Declaration of Conflicting Interest

The authors declare that they have no conflicts of interest, known competing financial interests, or personal relationships that could have influenced the work reported in this paper.

REFERENCES

- [1] M. U. Akmal, S. George Mathias, S. Asif, L. Koval, S. Knollmeyer, and D. Großmann, "Humanizing ai: A human-centered architecture to developing trustworthy intelligent systems," *Computer Science & Information Technology (CS & IT)*, 2025.
- [2] A. Fenwick and G. Molnar, "The importance of humanizing ai: using a behavioral lens to bridge the gaps between humans and machines," *Discover Artificial Intelligence*, vol. 2, no. 1, p. 14, 2022.
- [3] M. Fahrurrozi, S. R. Djalilah, D. Yunitasari, M. Mohzana, and H. S. Butt, "Development of technopreneurship based e-modules on entrepreneurial learning," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 3, pp. 795–808, 2026.
- [4] A. Fenwick, G. Molnar, and P. Frangos, "The critical role of hrm in ai-driven digital transformation: a paradigm shift to enable firms to move from ai implementation to human-centric adoption," *Discover Artificial Intelligence*, vol. 4, no. 1, p. 34, 2024.
- [5] D. Singh, "Advanced human-ai synergy for effective collaboration: ethical implications and future prospects," *INTELETICA*, vol. 3, no. 5, pp. 24–38, 2026.
- [6] S. L. Sitorus, A. Hermawan, A. W. Widjaja, and M. P. Berlianto, "Enhancing digital fashionpreneurship through innovation capability and market sensing," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 2, pp. 645–658, 2026.
- [7] F. Branda, M. Ciccozzi, and F. Scarpa, "Artificial intelligence in scientific research: Challenges, opportunities and the imperative of a human-centric synergy," *Journal of Informetrics*, vol. 19, no. 4, p. 101727, 2025.
- [8] A. Rizky, U. Rahardja, M. Muhtarom, D. N. Ramadhan, S. Triandari, and R. F. Terizla, "Responsible ai governance in decentralized web systems," *Legal Autonomous Web-based Governance (LAW)*, vol. 1, no. 1, pp. 123–133, 2026.
- [9] U. A. Usmani, A. Happonen, and J. Watada, "Human-centered artificial intelligence: Designing for user empowerment and ethical considerations," in *2023 5th international congress on human-computer interaction, optimization and robotic applications (HORA)*. IEEE, 2023.
- [10] I. N. Pratiwi, D. D. O. Prabawati, E. D. Wahyuni, N. Nursalam, I. Y. Widyawati, and N. A. Yahaya, "Entrepreneurship in social media literacy and intentions for diabetes prevention among adolescent students," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 1, pp. 85–98, 2026.
- [11] S. Banerjee and S. Griffiths, "Involving patients in artificial intelligence research to build trustworthy systems," *AI & SOCIETY*, vol. 39, no. 6, pp. 3037–3039, 2024.
- [12] A. H. Poole, "Data flourishing: Developing human-centered data science through communities of ethical practice," *Proceedings of the association for information Science and Technology*, vol. 60, no. 1, pp. 338–352, 2023.
- [13] A. Rozi, J. Junengsih, S. Alam, A. Pawar, W. Sumarjo, and D. Sunarsi, "Impact of hr management on ai implementation and data protection in indonesian manufacturing," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 6, no. 1, pp. 51–64, 2026.

- [14] H. Hamdan, U. Rahardja, M. Muhtarom, and A. Valerry, "Transforming education management with orange servant leadership and happiness practices," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 2, pp. 741–753, 2026.
- [15] R. Fioravante and A. Vaccaro, *Humanism and Artificial Intelligence*. Springer, 2025.
- [16] S. Sunanto, K. Kasmad, and J. Andriani, "Predicting professional entrepreneurial intention through core determinants of entrepreneurial attitude," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 6, no. 1, pp. 25–36, 2026.
- [17] J. R. Schoenherr, R. Abbas, K. Michael, P. Rivas, and T. D. Anderson, "Designing ai using a human-centered approach: Explainability and accuracy toward trustworthiness," *IEEE Transactions on Technology and Society*, vol. 4, no. 1, pp. 9–23, 2023.
- [18] J. R. Schoenherr, *Ethical artificial intelligence from popular to cognitive science: trust in the age of entanglement*. Routledge, 2022.
- [19] U. Rahardja, G. P. Cesna, R. Nuraeni, M. S. Gunawan, and J. Parker, "An analysis of the orange technology approach in short video-based digital marketing strategies," *ADI Bisnis Digital Interdisiplin Jurnal*, vol. 7, no. 1, pp. 15–24, 2026.
- [20] M. Siahaan, S. Kosasi, N. Sukendri, and A. Husain, "Enhancing smes business performance through strategic digital transformation," *IAIC Transactions on Sustainable Digital Innovation (ITSIDI)*, vol. 7, no. 1, pp. 85–96, 2025.
- [21] T. R. Merlo and J. Liebowitz, *Ethical AI and Data Science: Building Trustworthy and Transparent Systems*. CRC Press, 2026.
- [22] M. Tahri Sqalli, B. Aslonov, M. Gafurov, and S. Nurmatov, "Humanizing ai in medical training: ethical framework for responsible design," *Frontiers in Artificial Intelligence*, vol. 6, p. 1189914, 2023.
- [23] T. L. Anita, R. G. Freedy, N. Silawati, and M. Sunengsih, "Mathematical logic as a foundation for ai-driven decision-making systems," *Smart Human-centered Emerging Research in Machine Intelligence*, vol. 2, no. 1, pp. 14–25, 2026.
- [24] N. Bienefeld, E. Keller, and G. Grote, "Human-ai teaming in critical care: a comparative analysis of data scientists' and clinicians' perspectives on ai augmentation and automation," *Journal of Medical Internet Research*, vol. 26, p. e50130, 2024.
- [25] K. D. Hartomo, M. Zaki, G. K. Hanum, N. Silawati, and A. Valerry, "Empirical studies on the relationship between wearable stress detection and workplace productivity," *Journal of Orange Technology*, vol. 1, no. 1, pp. 1–10, 2024.
- [26] K. Achuthan, S. Ramanathan, S. Srinivas, and R. Raman, "Advancing cybersecurity and privacy with artificial intelligence: current trends and future research directions," *Frontiers in big data*, vol. 7, p. 1497535, 2024.
- [27] M. Cecconi, M. Greco, B. Shickel, D. C. Angus, H. Bailey, E. Bignami, T. Calandra, L. A. Celi, S. Einav, P. Elbers *et al.*, "Implementing artificial intelligence in critical care medicine: a consensus of 22," *Critical care*, vol. 29, no. 1, p. 290, 2025.
- [28] I. Sembiring, Q. Aini, R. A. Sunarjo, D. Julianingsih, and M. Rodriguez, "Blockchain enabled emotional wellbeing data sharing platform with humanistic ethical governance model," *Journal of Orange Technology*, vol. 1, no. 2, pp. 91–102, 2025.
- [29] P.-Z. Hahne and A. Schmoelz, "Trusting the machine: a digital humanist perspective on misplaced trust in artificial intelligence," *AI and Ethics*, vol. 6, no. 1, p. 115, 2026.
- [30] S. R. Bauskar, C. R. Madhavaram, E. P. Galla, J. R. Sunkara, and H. K. Gollangi, "Ai-driven phishing email detection: Leveraging big data analytics for enhanced cybersecurity," *Library Progress International*, vol. 44, no. 3, pp. 7211–7224, 2024.
- [31] A. Ruangkanjanases, N. Ivanov, M. Muhtarom, M. Salmi, and N. Lutfiani, "Advancing preventive community medicine for enhancing public health awareness," *Health, Empathy, and AI Learning (HEAL)*, vol. 1, no. 1, pp. 67–77, 2025.
- [32] D. Tuia, K. Schindler, B. Demir, X. X. Zhu, M. Kochupillai, S. Džeroski, J. N. Van Rijn, H. H. Hoos, F. Del Frate, M. Datcu *et al.*, "Artificial intelligence to advance earth observation: A review of models, recent trends, and pathways forward," *IEEE Geoscience and Remote Sensing Magazine*, 2024.
- [33] M. T. Sqalli, B. Aslonov, M. Gafurov, and S. Nurmatov, "Humanizing ai in medical training: Ethical framework for responsible design," *Frontiers in Artificial Intelligence*, vol. 6, p. 1189914, 2023. [Online]. Available: <https://doi.org/10.3389/frai.2023.1189914>