JSRM
JOURNAL OF SCHOLARLY RESEARCH AND METRICS

RESEARCH ARTICLE
Vol. 1, No. 1 (2026), pp. 27–44
Published online: 30 September 2026
*Open Access · CC BY 4.0*
**DOI:** https://doi.org/10.68708/jsrm.v1i1.03

# Mapping the Evolution of Human-in-the-Loop Artificial Intelligence: A Bibliometric Analysis

**Asif Altaf**
Systems Librarian, University of Applied Sciences Mittweida, Mittweida, Germany
Email: asifaltaf1987@gmail.com | ORCID: https://orcid.org/0000-0001-6106-6124

*Received: 05 September 2026 | Revised: 18 September 2026 | Revised: 19 September 2026 | Accepted: 26 September 2026 | Published online: 30 September 2026*

**How to cite:** Altaf, A. (2026). Mapping the evolution of human-in-the-loop artificial intelligence: A bibliometric analysis. *Journal of Scholarly Research and Metrics, 1*(1), 27–44. https://doi.org/10.68708/jsrm.v1i1.03

## Abstract

**Context:** Human-in-the-loop (HITL) artificial intelligence has evolved from human-assisted model training toward broader human–AI collaboration, oversight, and decision-making. Its rapidly expanding and multidisciplinary literature, however, remains fragmented across technological and application domains, creating a need for systematic mapping of the field.

**Aim:** This study maps the evolution, scholarly impact, geographical and institutional structure, collaboration patterns, and conceptual development of HITL AI research from 2016 to 2026.

**Approach:** A bibliometric analysis was conducted on 2,071 Scopus-indexed publications. Bibliometrix and its Biblioshiny interface were used for bibliometric performance analysis, and VOSviewer was used for keyword co-occurrence mapping and visualization.

**Key Findings:** HITL AI research recorded a 55.9% annual growth rate and involved 8,319 authors publishing across 1,029 sources. The United States and China led publication output. Keyword analysis identified 462 keywords organized into five clusters, with prominent themes including deep learning, reinforcement learning, generative AI, decision support, responsible AI, ethics, and human oversight.

**Interpretation:** The findings indicate that HITL AI is developing into an increasingly multidisciplinary field in which human participation extends beyond model training toward collaboration, oversight, and accountable AI deployment.

**Contribution:** This study provides an integrated bibliometric mapping of HITL AI that consolidates its growth, scholarly impact, leading contributors, collaboration patterns, and conceptual structure while identifying emerging directions for future research.

**Keywords:** human-in-the-loop, artificial intelligence, human–AI collaboration, bibliometric analysis, human oversight, responsible AI

## 1. Introduction

Artificial Intelligence (AI) has come a long way from rule-based and task-specific automation to machine-learning systems, which are capable of making more complex decisions. AI, even fully automated, still has its limits in contexts that demand context, in-depth knowledge, ethics, understanding and responsibility. The restrictions are especially significant in critical applications like healthcare, finance, legal decisions, and autonomous systems (Lazaros et al., 2026). Human-in-the-loop (HITL) AI has, therefore, become a strategy that intentionally leverages the human component for various phases of the AI lifecycle, ranging from design and training to deployment and maintenance.

The conceptual scope of HITL has broadened considerably as machine learning has advanced. Holzinger (2016) defined interactive machine learning as machine learning algorithms

© 2026 The Author. Published by ScholarVista Private Limited under the CC BY 4.0 licence (https://creativecommons.org/licenses/by/4.0/). 

that work with agents, such as humans, and adapt their learning behavior based on their interactions with agents, especially when the number of samples is small, the occurrence of events is rare, or the problem is too complex to be efficiently handled by purely algorithmic methods. Later research extended HITL to the stage of training a model with an intervention and to the level of human involvement in a system. For instance, Wu et al. (2022) categorized the HITL approaches into three subclasses: data processing, interventional model training, and independent system-level HITL design. This development puts people more and more into the role of co-worker than of labeler or end-stage supervisor.

This collaboration has meaning in all application fields. In the medical field, HITL combines clinical knowledge and machine intelligence to tackle issues of safety, accountability, and interpretability (Olawade et al., 2026). Human operators can supervise partially autonomous systems and step in when things go wrong; Liu, Nasiriany, et al. (2025) showed that this cooperation between humans and robotics can be used to help deploy and learn ongoing. HITL has also branched out into the human-centric manufacturing sector (Chen et al., 2025) and education (Memarian & Doleck, 2024). The developments are a sign that HITL is emerging as a multi-disciplinary paradigm that concerns not only model performance but also meaningful human judgment, collaboration, and oversight of the models.

An important conceptual distinction concerns human-in-the-loop (HITL) and human-on-the-loop configurations. In HITL systems, human input is incorporated directly into one or more stages of the computational, learning, or decision-making process, for example through data labeling, feedback, validation, correction, model refinement, or decision intervention (Holzinger, 2016; Lazaros et al., 2026; Mosqueira-Rey et al., 2023; Wu et al., 2022). Human participation can therefore form an integral part of the learning or decision cycle rather than being limited to passive observation. Depending on the application, humans may provide training examples, correct model outputs, guide learning, validate uncertain cases, or contribute domain expertise to improve system performance (Chen et al., 2025; Retzlaff et al., 2024; Wu et al., 2022). Human-on-the-loop configurations, in contrast, generally describe a more supervisory form of human involvement, in which the automated system operates with greater autonomy while the human monitors its behavior and retains the capacity to intervene, override, or assume control when necessary (Lazaros et al., 2026). Thus, as the operational autonomy of an AI system increases, the human role may shift from continuous participation within the computational loop toward monitoring and exception-based intervention. Nevertheless, these categories should not be treated as universally fixed or mutually exclusive because the nature and degree of human involvement vary across learning stages, deployment settings, application domains, and levels of automation (Lazaros et al., 2026; Mosqueira-Rey et al., 2023; Retzlaff et al., 2024). Accordingly, the term “human-on-the-loop” was included in the present search strategy to capture literature addressing this closely related supervisory form of human oversight, while HITL AI remained the principal conceptual focus of the bibliometric analysis.

However, the knowledge base is still siloed in multiple domains, including machine learning, human–computer interaction, explainable AI, healthcare, autonomous systems, governance and other application-specific areas. Current reviews have added much to the syntheses on HITL concepts, methods, taxonomies, and applications (Lazaros et al., 2026), but they have failed to provide large-scale quantitative mapping of publication growth, citation impact, productive sources, institutional involvement, international collaboration, and conceptual relationships. This is especially significant because of the recent growth and diversification in the field.

Accordingly, the present study conducts a bibliometric analysis of 2,071 Scopus-indexed HITL AI publications (2016–2026) to map the evolution and knowledge structure of the field. It investigates temporal trends in publications and citations, key publication sources, leading institutions and countries, international collaboration, and concepts identified through keyword

co-occurrence. In doing so, the study provides an integrated quantitative view of the development of HITL AI and the focus of emerging studies.

### 1.1 Research Objectives and Questions

This research aims to systematically identify the development, academic contributions, spatial and institutional landscape, cooperation trend, and conceptual development of HITL AI research from 2016 to 2026. Four research questions are explored:

RQ1: What is the trend of the volume and citation impact of HITL AI studies from 2016 to 2026?

RQ2: What are the main scholarly journals and publications for HITL AI research?

RQ3: What are the main institutions and countries for these contributions, and what are the trends in international research cooperation?

RQ4: What are the key ideas and relationships within the intellectual framework of the HITL AI research and what new areas of study are emerging that will shape the future direction of AI research?

## 2. Literature Review

### 2.1 Conceptual Foundations and Evolution of Human-in-the-Loop Artificial Intelligence

The conceptual roots reach across fields of cybernetics, cognitive science, decision theory and human-factors engineering, and have always been linked to the desire to marry computational power with human judgment (Lazaros et al., 2026). Feedback-oriented human–machine systems were actually developed before the advent of modern AI and continue to shape decision-support systems where computational systems supplement but do not supplant the human decision maker (Lazaros et al., 2026).

As machine learning developed this relationship shifted from human oversight of machine-produced results to human involvement in the learning process. Holzinger (2016) defined interactive machine learning as algorithms that can interact with agents (including humans) and that can optimize their learning behavior based on the interaction. This method is especially useful in cases where data is scarce, events are rare, uncertainty exists, and/or computational complexity is an issue in fully-automated learning. This was further realized in the active learning approach which allowed algorithms to ask human annotators for informative labels selectively, which led to less labeling and more focus on human effort where it could contribute the most to the learning process (Toumanidis et al., 2021).

As research in HITL progressed, the role of the human being became more and more distinct. Mosqueira-Rey et al. (2023) have adopted an active learning approach whereby the algorithm plays the main role in the learning process, an interactive machine learning approach that requires frequent and incremental human interaction, and a machine teaching approach where domain experts have more control over the knowledge transferred to the model. HITL, as a result, matured from data annotation to training, validating, explaining, correcting, and supporting the decision-making process of the model. Similarly, Wu et al. (2022) structured HITL development in terms of data processing, intervention in the learning process of the model, and system-level design of the HITL, showing the gradual escalation of human interventions throughout the machine-learning lifecycle.

More recently, HITL has expanded toward human–AI collaboration and hybrid intelligence with a focus on complementary abilities as opposed to substitutive abilities. Feedback, guidance, monitoring, and adaptation are processes that can be continuous if humans are involved, dependent on users' goals and contexts (Gómez-Carmona et al., 2024). One extension is reinforcement learning from human feedback, which adds human preferences to iterative agent learning (Retzlaff et al., 2024). With the advent of generative AI and large language models, the nature of HITL has been further redefined toward verification, contextual correction, trust, and responsible oversight (Hanafi et al., 2022; Lazaros et al., 2026).

## 2.2 Human–AI Collaboration, Oversight, and Application Domains

HITL artificial intelligence is reimagining AI not as a replacement for human skills, but part of a decision process. Human intervention may be a form of feedback, validation, intervention, interpretation or final decision and the level of intervention varies depending upon the task, the risk level and the time of intervention. For instance, Mosqueira-Rey et al. (2023) identified three HITL approaches: active learning, interactive machine learning, and machine teaching. Other more recent syntheses place human monitoring, feedback, and decision making within different parts of the AI process, and suggest healthcare, autonomous systems, cyber security, finance, education, and manufacturing as relevant application areas (Lazaros et al., 2026).

One of the core aspects of this cooperation is to maintain sense of context and accountability in the face of algorithmic outputs with ramifications. In health care, Salloch and Eriksen (2024) claimed that meaningful human involvement should not be limited to clicking an algorithmic recommendation. They considered the patients as “epistemic partners” or “fellow workers” in understanding the output from the machine, shifting the focus from supervision to co-reasoning in machine-learning based clinical decision support. This is especially important since the current clinical application of HITL in healthcare covers diagnostic imaging, clinical decision support, drug discovery, patient monitoring, and research, where algorithms were used in conjunction with the professional review and judgment by experts (Olawade et al., 2026). Human involvement is therefore not only intended to give the final approval but to add knowledge and context to the model that might not be captured in the model itself.

Even in the case of autonomous and robotic systems, the critical role of human involvement is obvious. Liu, Nasiriany, et al. (2025) introduced a division-of-work model that lets robots execute tasks in which they can do the job reliably, and human operators observe the action and intervene when it is hard. In simulation, they improved the rate of policy success by 8% compared to their reported state-of-the-art baselines, and by 27% in physical hardware, while also requiring less memory usage and converging faster. Human knowledge can also solve problems in robot training efficiency, real-time feedback, and adaptation to new and unexpected situations in smart manufacturing. To articulate HITL robot learning, Chen et al. (2025) discussed the perception, cognition, behavior, and empathy of humans and explored their potential applications in human-centric manufacturing environments. Another industrial use case is predictive maintenance, as Amaliah et al. (2025) highlighted transparency and trustworthiness as key challenges for AI adoption and the need for alignment of three domains: HITL explainable AI (XAI) systems, operation workflows, and the cognitive needs of maintenance workers.

Human oversight should not, however, be regarded as an automatic safeguard. In certain scenarios, like in the police, social worker, or online content moderation processes, hybrid human–algorithmic decision-making is often added in order to preserve agency and accountability, offer legal protection, and facilitate quality control, as discussed by Enarsson et al. (2022). However, empirical evidence shows there is a key limitation. In an online experimental study with 292 participants, Sele and Chugunova (2024) showed that the participants' ability to monitor and adjust algorithmic recommendations led to a 7-point increase in the percentage of those who preferred algorithmic delegation but that overall decision accuracy decreased in the experiment. This finding is a reminder that human presence alone does not guarantee effective oversight.

The literature suggests that the key to the value of HITL is in the design of the human–AI collaboration, which must be meaningful and context-appropriate, rather than merely putting a person into an automated process. Good oversight needs to involve clear decisions about who has the authority to make decisions, the understandable output of AI, opportunities for inputting or overriding the AI, calibrated levels of trust, and clear accountability. Requirements differ by domain, such as real-time intervention, which is essential in autonomous systems (Liu, Nasiriany, et al., 2025); agency and fairness, which are key in legal and administrative applications (Enarsson et al., 2022); and quality inspection and adaptation in manufacturing (Chen et al., 2025); and clinical validation and judgment in healthcare (Olawade et al., 2026). From the cross-domain

research, Amaliah et al. (2025) and Lazaros et al. (2026) point out that the hurdles for effective HITL deployment persist in explainability, cognitive load, trust calibration, scalability, regulatory compliance, and value alignment.

Although existing studies have substantially advanced understanding of HITL concepts, methods, and domain-specific applications, the literature remains fragmented across different technological and application areas. Previous reviews have primarily examined particular HITL methods or specific contexts, such as education, healthcare, human–AI collaboration, and hybrid intelligence (Memarian & Doleck, 2024; Liu, Tang, et al., 2025; Lazaros et al., 2026; Olawade et al., 2026). Consequently, there remains limited large-scale quantitative evidence showing how HITL AI has developed as an integrated research field in terms of publication growth, citation impact, leading sources and institutions, geographical contributions, international collaboration, and conceptual structure. The present study addresses this gap through a bibliometric mapping of Scopus-indexed HITL AI research from 2016 to 2026.

## 3. Method

### 3.1 Research Design

The research design used in this study was a bibliometric and science-mapping research design. Analytical issues included the growth of publications, impact of citations, core sources, institutional and geographical contributions, international collaboration, and conceptual structure of the field.

### 3.2 Data Source

The bibliographic information was available from the Scopus database. For analyses of publications, citations, affiliations, countries and keywords, the use of structured bibliographic metadata was required - hence the selection of Scopus. A search and final data extraction was carried out on 26 August 2026.

### 3.3 Search Strategy

The search strategy combined HITL terminology, principally human-in-the-loop, humans-in-the-loop, and human-on-the-loop, with AI-related terms covering artificial intelligence, machine learning, deep learning, active learning, reinforcement learning, large language models, generative AI, human oversight, human–AI interaction, decision support, and annotation. Engineering-specific in-the-loop terminology and exoskeleton-related terms were excluded to improve topical relevance.

( TITLE ( "human-in-the-loop" OR "humans-in-the-loop" OR "human-on-the-loop" ) OR KEY ( "human-in-the-loop" OR "humans-in-the-loop" OR "human-on-the-loop" ) OR TITLE-ABS-KEY ( "human-in-the-loop" OR "humans-in-the-loop" OR "human-on-the-loop" ) AND TITLE-ABS-KEY ( "artificial intelligence" OR "machine learning" OR "deep learning" OR "active learning" OR "reinforcement learning" OR "large language model*" OR "generative AI" OR "human oversight" OR "human-AI" OR "decision support" OR annotat* ) AND NOT TITLE-ABS-KEY ( "hardware-in-the-loop" OR "software-in-the-loop" OR "power-hardware-in-the-loop" OR "vehicle-in-the-loop" OR "controller-in-the-loop" OR "model-in-the-loop" OR "processor-in-the-loop" OR "chip-in-the-loop" OR "battery-in-the-loop" ) AND NOT TITLE-ABS-KEY ( exoskeleton* OR exosuit* OR orthos* OR prosthe* OR "metabolic cost" OR "gait" OR "walking assistance" OR "ankle exo*" ) ) AND ( EXCLUDE ( SUBJAREA , "MULT" ) OR EXCLUDE ( SUBJAREA , "MATH" ) OR EXCLUDE ( SUBJAREA , "DENT" ) OR EXCLUDE ( SUBJAREA , "IMMU" ) OR EXCLUDE ( SUBJAREA , "VETE" ) ) AND ( EXCLUDE ( DOCTYPE , "cp" ) OR EXCLUDE ( DOCTYPE , "tb" ) OR EXCLUDE ( DOCTYPE , "le" ) OR EXCLUDE ( DOCTYPE , "no" ) OR EXCLUDE ( DOCTYPE , "bk" ) OR EXCLUDE ( DOCTYPE , "ed" ) OR EXCLUDE ( DOCTYPE , "cr" ) OR EXCLUDE ( DOCTYPE , "ch" ) )

 DOI: https://doi.org/10.68708/jsrm.v1i1.03 

AND ( EXCLUDE ( LANGUAGE , "French" ) OR EXCLUDE ( LANGUAGE , "Malay" ) OR EXCLUDE ( LANGUAGE , "Turkish" ) OR EXCLUDE ( LANGUAGE , "Ukrainian" ) OR EXCLUDE ( LANGUAGE , "Chinese" ) OR EXCLUDE ( LANGUAGE , "German" ) OR EXCLUDE ( LANGUAGE , "Spanish" ) OR EXCLUDE ( LANGUAGE , "Italian" ) OR EXCLUDE ( LANGUAGE , "Japanese" ) OR EXCLUDE ( LANGUAGE , "Persian" ) ) AND PUBYEAR > 2015 AND PUBYEAR < 2027

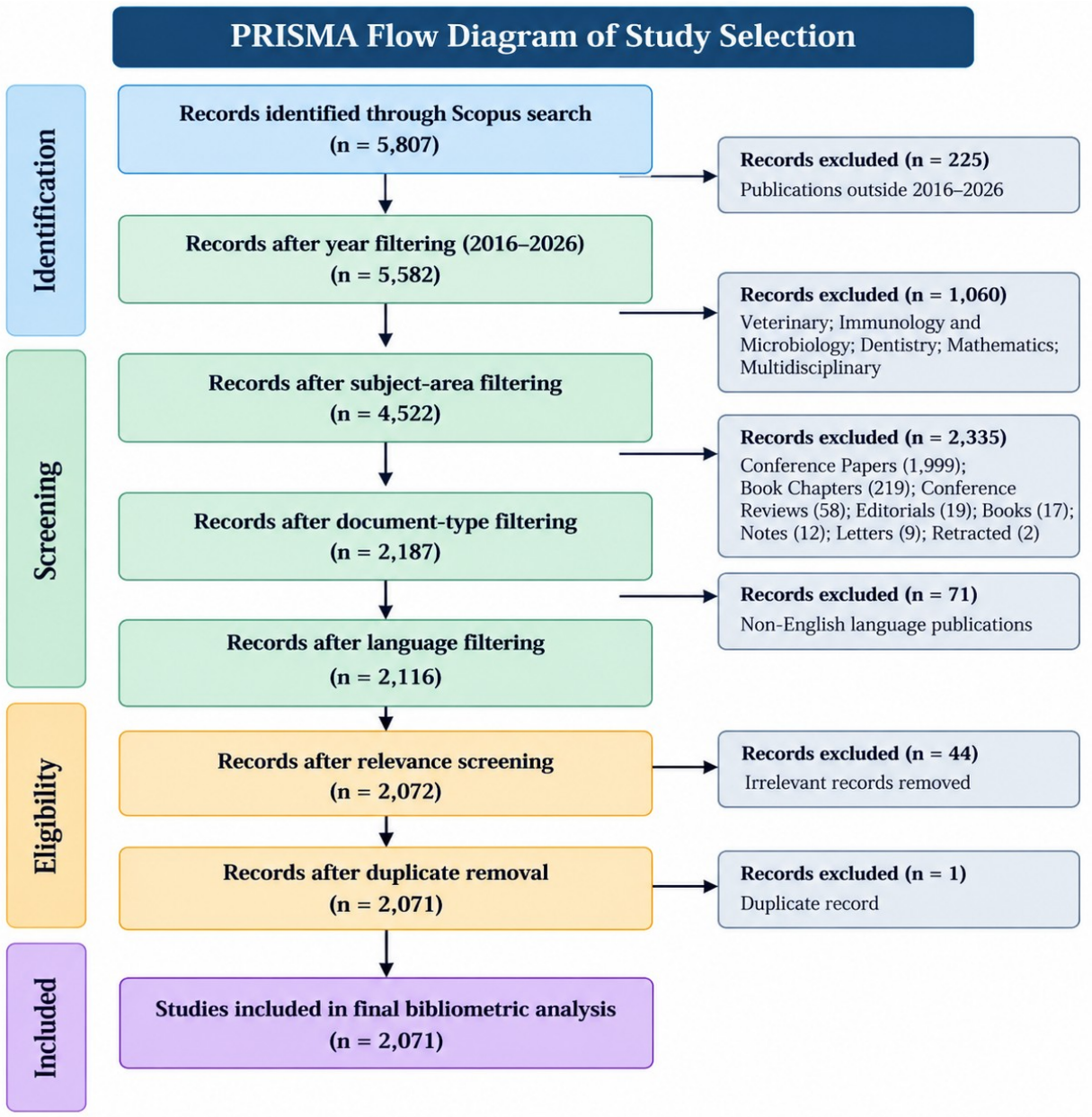
The year 2016 was selected as the starting point to capture the contemporary development of human-in-the-loop artificial intelligence while maintaining a sufficiently focused and comparable observation window. This boundary does not imply that human involvement in computational or automated systems originated in 2016; related ideas have substantially earlier roots in cybernetics, human factors, decision support, and interactive computing. Rather, 2016 represents a useful point of departure for the modern HITL-AI literature because influential work during this period explicitly articulated human-in-the-loop and interactive machine-learning approaches in relation to contemporary machine-learning systems (e.g., Holzinger, 2016). The 2016–2026 window was therefore adopted to examine the recent evolution of HITL within the modern AI and machine-learning landscape.

### 3.4 Eligibility Criteria

Publications were included if they were published between 2016 and 2026, written in English, and classified as articles, reviews, or short surveys. Records assigned to the Veterinary; Immunology and Microbiology; Dentistry; Mathematics; and Multidisciplinary subject categories were excluded. Conference papers, book chapters, conference reviews, editorials, books, notes, letters, and retracted records were also removed.

### 3.5 Study Selection

Figure 1. PRISMA-Based Study Selection and Screening Process




Source: Author's compilation based on Scopus data.

Figure 1 shows the sequential screening and selection procedure used to identify the final set of publications. The initial Scopus search retrieved 5,807 records. Restricting the publication period to 2016–2026 excluded 225 records, leaving 5,582. Subject-area filtering removed 1,060 records, leaving 4,522. Document-type filtering excluded 2,335 records, leaving 2,187. Restriction to English-language publications removed a further 71 records, leaving 2,116. Relevance screening excluded 44 records, leaving 2,072, and the removal of one duplicate produced a final dataset of 2,071 publications. In total, 3,736 records were excluded across the complete screening process.

Following the database-level eligibility filters, the remaining records underwent an additional record-level screening before bibliometric analysis. The bibliographic metadata of each potentially problematic record were examined using the publication title, DOI, author information, author/indexed keywords, and abstract, where available. Records were excluded when they contained insufficient bibliographic information for reliable identification and analysis, including missing author information, missing keywords, or missing abstracts, and when the available title, DOI, and other metadata did not provide sufficient information to establish relevance to the scope

of human-in-the-loop artificial intelligence. Records identified through the title, abstract, keywords, and DOI as unrelated to the conceptual scope of HITL AI were also removed.

This record-level screening resulted in the exclusion of 44 publications. The screening was therefore not based solely on keyword matching; rather, the available bibliographic fields were examined collectively to determine whether each record contained adequate metadata and represented substantive research on HITL AI. Following this screening, 2,072 records remained. One duplicate record was subsequently identified and removed, producing the final dataset of 2,071 publications used for bibliometric analysis.

### 3.6 Data Cleaning and Harmonization

Before analysis, bibliographic metadata were checked for duplicate and irrelevant records. For the keyword-network analysis, synonymous and variant expressions for example, variants of HITL, artificial intelligence, machine learning, and large language models were harmonized, and generic indexing terms were removed to reduce noise.

### 3.7 Bibliometric Performance Analysis

Descriptive and performance indicators included annual scientific production, annual growth rate, citation impact, source productivity, authorship and collaboration indicators, institutional productivity, and corresponding-author countries. Bradford's law was applied to identify the core publication sources.

### 3.8 Science Mapping Analysis

Science mapping was conducted through keyword co-occurrence analysis in order to reveal the conceptual structure and major thematic relationships within HITL AI research.

### 3.9 Analytical Software and Network Construction

Bibliometric performance indicators were generated using the Bibliometrix R package and its Biblioshiny interface (Aria & Cuccurullo, 2017). The final bibliographic dataset was exported directly from Scopus using the standard bibliographic fields available through the Scopus export function and was subsequently used for bibliometric performance and science-mapping analyses.

VOSviewer (van Eck & Waltman, 2010) was used to construct and visualize the keyword co-occurrence network. Keyword data available in the Scopus records were used as the basis for the analysis. Before network construction, synonymous expressions, spelling variants, abbreviations, and closely related forms of key concepts were harmonized, while generic or non-informative indexing terms were removed to reduce noise in the network. This cleaning process included variants associated with human-in-the-loop, artificial intelligence, machine learning, and large language models, as described in Section 3.6.

The cleaned keyword data were then analyzed in VOSviewer using its standard network-construction and visualization procedures. The resulting network statistics (numbers of keywords, clusters, links, and total link strength) were obtained from the VOSviewer output and were used to characterize the conceptual structure and relationships among the major themes in HITL AI research.

### 3.10 Reporting and Reproducibility

The identification, filtering, screening, and inclusion process was documented using a PRISMA-based flow diagram in order to support transparency and reproducibility. The final dataset comprised 2,071 publications from 1,029 sources and 8,319 authors, and provided the evidence base for the analyses reported below.

## 4. Results

### 4.1 Descriptive Characteristics of the Corpus

Table 1. Descriptive Characteristics of the Bibliometric Dataset on Human-in-the-Loop Artificial Intelligence Research

| Category | Indicator | Value |
|---|---|---|
| Main information about data | Timespan | 2016–2026 |
| | Sources (journals, books, etc.) | 1,029 |
| | Documents | 2,071 |
| | Annual growth rate (%) | 55.9 |
| | Document average age (years) | 1.54 |
| | Average citations per document | 15.71 |
| | References | 233,651 |
| Document contents | Keywords Plus (ID) | 11,135 |
| | Author's Keywords (DE) | 6,507 |
| Authors | Authors | 8,319 |
| | Authors of single-authored documents | 192 |
| Author collaboration | Single-authored documents | 197 |
| | Co-authors per document | 4.94 |
| | International co-authorships (%) | 27.38 |
| Document types | Articles | 1,840 |
| | Reviews | 225 |
| | Short surveys | 6 |
| | Total | 2,071 |

Source: Author’s analysis of Scopus data using Bibliometrix/Biblioshiny.

Table 1 summarizes the principal characteristics of the 2,071 publications included in the analysis. The literature expanded substantially, with an annual growth rate of 55.90% and a low average document age of 1.54 years, indicating a rapidly developing research domain. The publications received an average of 15.71 citations per document and collectively cited 233,651 references. Research activity was highly collaborative, involving 8,319 authors and an average of 4.94 co-authors per document, and international collaborations accounted for 27.38% of publications. Articles constituted the dominant document type (n = 1,840; 88.85%), followed by reviews (n = 225; 10.86%) and short surveys (n = 6; 0.29%).

### 4.2 Annual Publication Output and Citation Impact

Table 2. Annual Citation Impact and Average Citations per Year

| Year | Mean total citations per article | Documents (N) | Mean citations per year | Citable years |
|---|---|---|---|---|
| 2016 | 101.55 | 11 | 9.23 | 11 |
| 2017 | 81 | 15 | 8.1 | 10 |
| 2018 | 59.8 | 20 | 6.64 | 9 |
| 2019 | 71.26 | 46 | 8.91 | 8 |

| | | | | |
|---|---|---|---|---|
| 2020 | 60.71 | 51 | 8.67 | 7 |
| 2021 | 68.71 | 65 | 11.45 | 6 |
| 2022 | 51.63 | 127 | 10.33 | 5 |
| 2023 | 29.25 | 153 | 7.31 | 4 |
| 2024 | 16.42 | 205 | 5.47 | 3 |
| 2025 | 6.53 | 445 | 3.27 | 2 |
| 2026 | 0.92 | 933 | 0.92 | 1 |

Source: Author's analysis of Scopus data using Bibliometrix/Biblioshiny.

Table 2 shows substantial growth in publication output, from 11 documents in 2016 to 933 in 2026, indicating a sharp recent expansion of HITL AI research. Earlier publications show higher cumulative citation impact: 2016 recorded the highest mean total citations per article (101.55), partly reflecting its longer citation window. When citation exposure is taken into account, 2021 achieved the highest mean annual citation rate (11.45), followed by 2022 (10.33). The lower citation values for 2025 and 2026 should therefore be interpreted cautiously, because these publications have had substantially less time to accumulate citations.

### 4.3 Core Publication Sources

Table 3. Core Publication Sources According to Bradford's Law

| Rank | Source/Journal | Documents (n) | Cumulative documents (n) | Bradford zone |
|---|---|---|---|---|
| 1 | IEEE Access | 74 | 74 | Zone 1 |
| 2 | Applied Sciences (Switzerland) | 37 | 111 | Zone 1 |
| 3 | Journal of Medical Internet Research | 24 | 135 | Zone 1 |
| 4 | Expert Systems with Applications | 23 | 158 | Zone 1 |
| 5 | Frontiers in Artificial Intelligence | 22 | 180 | Zone 1 |
| 6 | Sensors | 21 | 201 | Zone 1 |
| 7 | Advanced Engineering Informatics | 19 | 220 | Zone 1 |
| 8 | Electronics (Switzerland) | 18 | 238 | Zone 1 |
| 9 | Engineering Applications of Artificial Intelligence | 16 | 254 | Zone 1 |
| 10 | Automation in Construction | 14 | 268 | Zone 1 |

Source: Author's analysis of Scopus data using Bibliometrix/Biblioshiny.

Table 3 identifies the leading core sources according to Bradford's law. IEEE Access is the most productive source, with 74 publications, substantially exceeding Applied Sciences (37). The Journal of Medical Internet Research (24) and Expert Systems with Applications (23) also contribute notably. Collectively, the 10 core sources listed account for 268 publications. Their disciplinary diversity, spanning artificial intelligence, engineering, healthcare, sensing

technologies, and construction, suggests that HITL AI has developed as a multidisciplinary research domain rather than one concentrated within a single application area.

### 4.4 Institutional and Geographical Contributions

Table 4. Most Productive Institutional Affiliations

| Rank | Institutional affiliation | Documents (n) |
|---|---|---|
| 1 | University of California | 41 |
| 2 | Beihang University | 30 |
| 3 | Tsinghua University | 30 |
| 4 | Nanyang Technological University | 28 |
| 5 | University of Toronto | 24 |
| 6 | Stanford University | 23 |
| 7 | Zhejiang University | 23 |
| 8 | Carnegie Mellon University | 20 |
| 9 | Northeastern University | 19 |
| 10 | The Hong Kong Polytechnic University | 19 |

Source: Author's analysis of Scopus data using Bibliometrix/Biblioshiny.

Table 4 presents the 10 most productive institutional affiliations. The University of California leads with 41 articles, followed by Beihang University and Tsinghua University with 30 articles each. Nanyang Technological University ranks fourth with 28 articles, and the University of Toronto contributes 24. Stanford University and Zhejiang University each account for 23 articles. The distribution of leading affiliations across North America and Asia indicates geographically diverse institutional engagement, with particularly strong contributions from universities in the United States, China, Canada, Singapore, and Hong Kong.

Table 5. Leading Countries by Corresponding Authorship and International Collaboration

| Rank | Country | Documents (n) | Documents (%) | SCP (n) | MCP (n) | MCP (%) |
|---|---|---|---|---|---|---|
| 1 | United States | 392 | 18.9 | 323 | 69 | 17.6 |
| 2 | China | 300 | 14.5 | 231 | 69 | 23 |
| 3 | Germany | 114 | 5.5 | 80 | 34 | 29.8 |
| 4 | United Kingdom | 104 | 5 | 58 | 46 | 44.2 |
| 5 | Italy | 75 | 3.6 | 54 | 21 | 28 |
| 6 | India | 71 | 3.4 | 54 | 17 | 23.9 |
| 7 | South Korea | 62 | 3 | 53 | 9 | 14.5 |
| 8 | Australia | 59 | 2.8 | 37 | 22 | 37.3 |

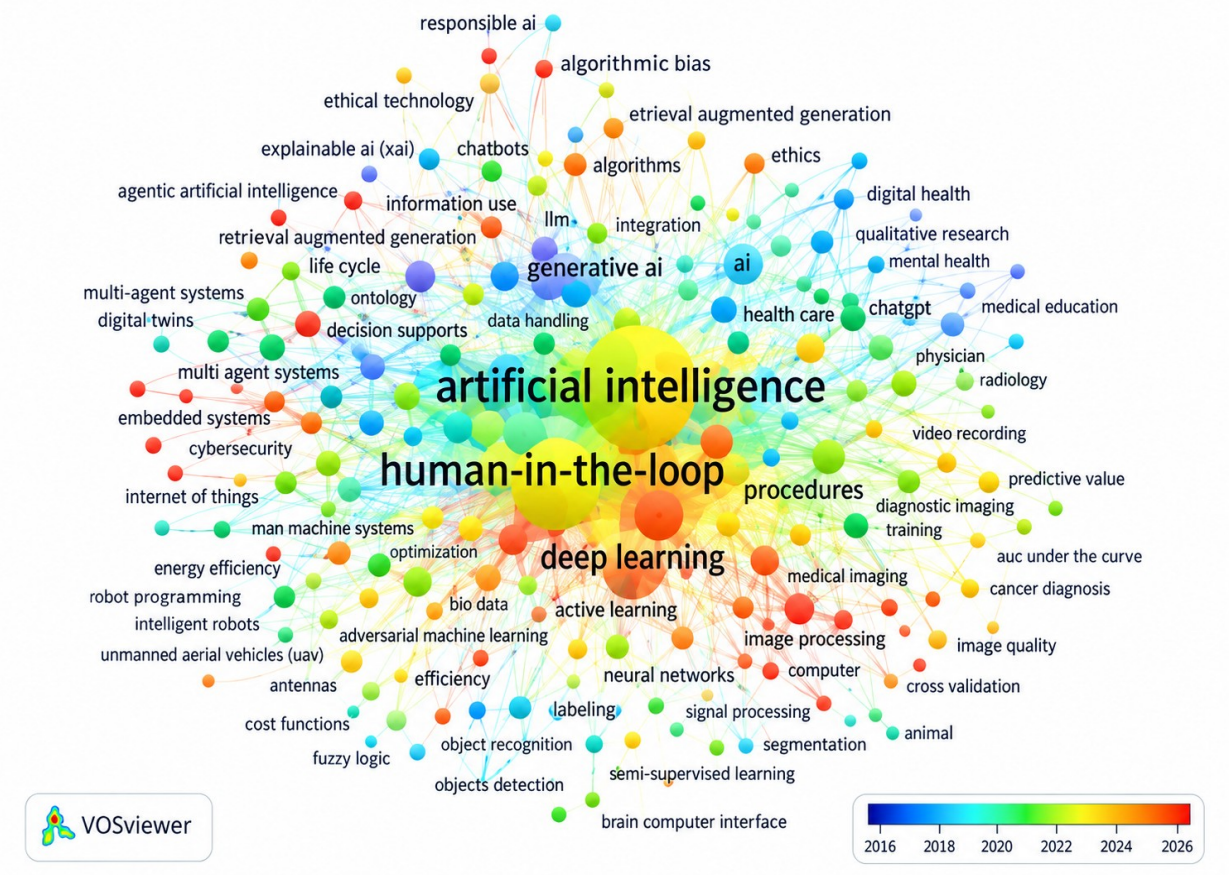
| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | Canada | 54 | 2.6 | 35 | 19 | 35.2 |
| 10 | Spain | 40 | 1.9 | 26 | 14 | 35 |

Note. SCP = single-country publications; MCP = multiple-country publications. Documents (%) is calculated on the full corpus of 2,071 publications. MCP (%) is the proportion of each country's publications involving international co-authorship. Countries are assigned by corresponding-author affiliation.
Source: Author's analysis of Scopus data using Bibliometrix/Biblioshiny.

Table 5 shows the geographical distribution of corresponding authors and international research collaboration. The collaboration results reveal two distinct dimensions of international research activity. In absolute terms, the United States and China each recorded 69 multiple-country publications (MCPs), representing the largest volumes of internationally co-authored output among the countries listed. However, international collaboration intensity differs when MCPs are considered relative to each country's total corresponding-author output. On this proportional measure, the United Kingdom recorded the highest MCP rate among the 10 leading countries (44.2%), followed by Australia (37.3%), Canada (35.2%), and Spain (35.0%). Thus, high publication or collaboration volume should not be interpreted as equivalent to a high propensity for international collaboration.

### 4.5 Conceptual Structure of the Field

Figure 2. Keyword Co-occurrence Network of Human-in-the-Loop Artificial Intelligence Research



Source: Author’s visualization based on Scopus data using VOSviewer.

Figure 2 presents the keyword co-occurrence network, which comprised 462 keywords organized into five clusters, with 26,643 links and a total link strength of 64,513. The network is strongly centered on human-in-the-loop and artificial intelligence, indicating their dominant and highly interconnected roles within the knowledge structure of the field. Other prominent themes include deep learning, reinforcement learning, active learning, generative AI, decision support, intelligent agents, and healthcare applications. The dense interconnections demonstrate substantial integration between machine-learning methodologies and human participation in AI development and decision-making. Emerging themes surrounding generative AI, retrieval-augmented generation, responsible AI, ethics, human oversight, and algorithmic bias further indicate a shift toward human-centered, accountable, and trustworthy AI systems. The network thus highlights the multidisciplinary and rapidly evolving character of HITL AI research.

## 5. Discussion

The results show that HITL AI is moving beyond a relatively narrow machine learning field towards a more general multidisciplinary research field focused on the integration of human decision-making into increasingly autonomous and sophisticated AI systems. The growing number of new publications and their spread across 1,029 sources indicate that HITL has moved beyond data annotation and model training. Instead, it is growing its research base to include interactive learning, decision support, human–AI collaboration, oversight, and responsible deployment. This interpretation aligns with prior conceptual research, which indicates that humans can be engaged in various phases of the machine-learning lifecycle, including data processing, model training, and interaction and decision-making in the system level (Lazaros et al., 2026; Mosqueira-Rey et al., 2023; Wu et al., 2022). As such, the bibliometric evidence suggests that the focus in HITL research is moving beyond a binary approach of "yes" to "no" and into a new conceptual paradigm that asks "how, when, and how much authority do humans have when participating in an AI system?".

The temporal pattern supports this interpretation, but requires careful interpretation. The number of publications rose significantly during the time period, especially in the most recent years. The 2026 count, however, only indexes publications until 26 August 2026 and is not to be regarded as a completed year's publication or compared directly with previous full publication years. Likewise, the smaller citation figures for 2025 and 2026 are not necessarily attributable to less influence in the scholarly world, but are more likely due to their low citation period. The above-mentioned result for the earlier cohorts, specifically 2021 and 2022, demonstrates that it is crucial to distinguish between cumulative impact of citations and the intensity of citations over time. In general, the current expansion in the literature is aligned with the many ongoing developments of HITL, such as the incorporation of reinforcement learning with human feedback, generative AI, large language models, and the need for responsible human oversight (Hanafi et al., 2022; Lazaros et al., 2026; Retzlaff et al., 2024). The bibliometric results do not provide evidence that these developments had a direct influence on the surge in publications, however, they are part of the conceptual landscape and hence are more and more gaining importance in today's research on HITL.

In addition, the distribution of publications over the core sources further underlines the multidisciplinary nature of the field. While IEEE Access is the most productive individual source, the central sources of publications span across artificial intelligence, engineering, healthcare, sensing technologies, and applied computational research. This dispersion is relevant because it indicates that HITL is not only a domain specific technique, but also a cross-cutting research paradigm that is being adopted to different modes of human–AI interactions. Previous studies also show significant differences in the way HITL is implemented in different applications, including healthcare, manufacturing, interactive machine learning, and other application fields (Chen et al.,

2025; Lazaros et al., 2026; Olawade et al., 2026). The source structure therefore not only shows disciplinary diversity, but also that there is no single standardized model of how humans participate in the technical system; that depends on the context of the application, the risk profile, and the type of decision.

The geographical results show another difference between the productivity of research and international cooperation. The United States and China are the top contributors to corresponding-author publications, as well as the absolute number of co-authored publications with other countries (69 each). Their absolute collaboration volume should not be interpreted as being equal to collaboration intensity, however. If the share of international collaboration is compared to the output of each country, the United Kingdom has the highest MCP rate (44.2%), followed by Australia (37.3%), Canada (35.2%) and Spain (35.0%). This distinction reflects two aspects of the international organization of HITL research: a large research system can have a large absolute number of international publications and a small absolute number of publications can have a high collaborative potential. Therefore, publication leadership, absolute international collaboration volume, and proportional collaboration intensity should be considered analytically independent, although related attributes of the field.

The keyword network further shows that the intellectual framework of HITL is expanding. Deep learning, reinforcement learning, active learning, decision support, and intelligent agents signal the methodology of AI, while the rise of generative AI, retrieval-augmented generation, responsible AI, ethics, human oversight, and algorithmic bias indicate greater awareness of the impact of more powerful AI systems. This is a trend similar to those observed in previous studies, where humans took on roles ranging from labeling and training signals to interpretation, validation, correction, collaboration, and oversight (Gómez-Carmona et al., 2024; Lazaros et al., 2026; Mosqueira-Rey et al., 2023). This is an important point because it indicates that, in today's HITL research, concepts which focus on technical performance are increasingly being deployed in a network alongside concepts grounded in human authority, accountability and responsible deployment.

Meanwhile, the growth of human oversight is not a sign that humans on the interface will definitely improve the quality of decisions or safety of AI processes. For example, Sele and Chugunova (2024) found that when participants are able to monitor and make adjustments to algorithmic recommendations, they are more willing to delegate decisions to algorithms but also less accurate overall. It is also important to note that meaningful human participation requires more than just giving the thumbs-up to machine-generated recommendations, as stressed by Salloch and Eriksen (2024). These findings help explain why the concept of responsible AI and other ethics, human oversight, and algorithmic bias concerns are now associated with HITL research. The challenge is now moving from the mere presence of humans to the creation of ways to participate in which humans have the right and the opportunity to have information, authority, expertise and to intervene.

This interpretation also sheds light on the connection between human-in-the-loop and human-on-the-loop methods. In HITL the human's role in learning, feedback, corrections, validation and decision making is more active, and in human-on-the-loop, the role of the human is more to supervise the autonomous operation of the systems with the ability to intervene if needed (Lazaros et al., 2026). It is important to note that these are not mutually exclusive or fixed categories, and the level of human involvement can vary when training, deploying, and implementing various levels of autonomy in systems (Lazaros et al., 2026; Mosqueira-Rey et al., 2023; Retzlaff et al., 2024). This evolution in the current keyword network towards oversight-related concepts can be seen as a part of the general shift from human involvement in the model development phase to human supervision and governance throughout the AI lifecycle.

Overall, this study provides a field-level quantitative overview of HITL AI that complements the existing reviews. The concepts, methods, applications, and specific fields of HITL have been studied previously, such as in the interactive machine learning (Mosqueira-Rey et

al., 2023), education (Memarian & Doleck, 2024), healthcare (Olawade et al., 2026), reinforcement learning (Retzlaff et al., 2024), and hybrid intelligence (Liu, Tang, et al., 2025). The value of the present study lies not in suggesting another taxonomy of the HITL techniques, but in demonstrating the place of the various research lines within a wider bibliometric framework. The study also reveals that by analyzing the co-occurrence of keywords, the number of references, the types of publications, the institutions publishing and the geographic locations, as well as the extent of global collaboration, it is evidenced that the discipline of HITL AI is now organized around the interplay between techniques for technical learning, human-AI interaction and responsible management.

### 5.1 Implications

The findings have theoretical, methodological and practical implications. Theoretically, the results suggest that HITL AI is becoming a more integral research field where machine learning is interwoven with human judgment, collaboration, supervision, and responsible AI. Human roles are shifting from annotation and model training to roles of governance and validation of increasingly autonomous systems, with generative AI, explainability, ethics and human oversight becoming key features. From a practical perspective, the findings and prior literature underscore the importance of substantive human engagement rather than merely nominal human supervision in high-impact AI applications. Previous studies highlight the importance of meaningful interaction, clear decision authority, interpretability, and calibrated trust in HITL settings (Salloch & Eriksen, 2024; Sele & Chugunova, 2024). The study also serves as a structured guide for researchers to the leading journals, institutions, countries, collaboration networks, and thematic groups, which could help them find publication opportunities, researchers to work with, or new research avenues.

### 5.2 Limitations

This study has several limitations. First, the analysis uses only Scopus, which means that publications indexed only in other databases, such as Web of Science, PubMed, IEEE Xplore and others were not included. Second, by limiting the data to only English-language articles, reviews, and short surveys, language and/or document-type bias may be introduced. In the case of computer science and AI, the exclusion of conference proceedings is especially noteworthy as it is an important conduit for dissemination; 1,999 conference papers were filtered out at the document-type level. Third, citation indicators disadvantage newer publications because they have had fewer years to accumulate citations, making direct comparisons with older publications difficult. Fourth, keyword harmonization and the elimination of generic terms inevitably involve analytical judgment and can affect the network structure. Fifth, institutional affiliation counts were not broken down by campus, potentially inflating counts for multi-campus university systems. Lastly, HITL AI is a relatively new concept and the bibliometric landscape is changing, so the 2026 data represent the database only as of the 26 August 2026 collection date, not a completed publication year.

### 5.3 Future Research Directions

Future studies could expand bibliometric coverage by combining Scopus with other databases and by exploring conference publications to gain a more comprehensive view of AI scholarship. Longitudinal analyses could examine whether HITL publications have recently increased in pace and the trajectory of the new themes after 2026. The keyword network highlights areas of potential interest such as generative AI, large language models (LLMs), retrieval-augmented generation, responsible AI, ethical considerations, human oversight, and algorithmic bias. Future empirical research also needs to go beyond the question of the presence of humans in AI processes, and explore when, where, and how humans add value to the process. Specific attention should be paid to the areas of calibration, automation bias, cognitive workload, accountability, explainability, and

appropriate allocation of decision authority (Enarsson et al., 2022; Sele & Chugunova, 2024). Across application domains of healthcare, autonomous systems, education, finance, and manufacturing, comparative studies could test whether effective HITL design principles transfer across applications.

## 6. Conclusion

This study performed a bibliometric mapping of HITL AI research for the period 2016–2026, based on 2,071 Scopus-indexed publications. The results show rapid growth, strong interdisciplinarity, and increasing research collaboration. The United States and China lead in publication output, while the share of internationally co-authored publications differs widely among the leading countries. The conceptual network places human-in-the-loop and AI at the heart of themes of machine learning, deep learning, reinforcement learning, decision support, generative AI, and responsible AI. HITL is transitioning from a process to add human input to model building to a new paradigm of human–AI collaboration, monitoring, and accountability. This study offers a summary of the development of the field, the knowledge landscape of AI-human interaction, its geographical spread, and the emerging themes, and can serve as a basis for future work related to effective, trustworthy and context-responsive human input into AI systems.

## Declarations

### Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### Conflict of Interest

The author declares no conflict of interest.

### Author Contributions (CRediT)

Asif Altaf: Conceptualization, Methodology, Data Curation, Formal Analysis, Visualization, Investigation, Writing – Original Draft, Writing – Review & Editing.

### Ethics Approval and Informed Consent

Not applicable. This study is based exclusively on bibliographic data obtained from Scopus and did not involve human participants or animals.

### Declaration of Generative AI Use

No generative AI tools were used in preparing this article.

### Data Availability

The bibliographic data analyzed in this study were retrieved from the Scopus database on 26 August 2026. The analyzed data may be made available by the author upon reasonable request, subject to Scopus licensing and database-access conditions.

### Acknowledgements

Not applicable.

## References

Amaliah, N. R., Tjahjono, B., & Palade, V. (2025). Human-in-the-loop XAI for predictive maintenance: A systematic review of interactive systems and their effectiveness in maintenance decision-making. *Electronics*, *14*(17), 3384. https://doi.org/10.3390/electronics14173384

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Chen, H., Li, S., Fan, J., Duan, A., Yang, C., Navarro-Alarcon, D., & Zheng, P. (2025). Human-in-the-loop robot learning for smart manufacturing: A human-centric perspective. *IEEE Transactions on Automation Science and Engineering*, *22*, 11062–11086. https://doi.org/10.1109/TASE.2025.3528051

Enarsson, T., Enqvist, L., & Naarttijärvi, M. (2022). Approaching the human in the loop – legal perspectives on hybrid human/algorithmic decision-making in three contexts. *Information & Communications Technology Law*, *31*(1), 123–153. https://doi.org/10.1080/13600834.2021.1958860

Gómez-Carmona, O., Casado-Mansilla, D., López-de-Ipiña, D., & García-Zubia, J. (2024). Human-in-the-loop machine learning: Reconceptualizing the role of the user in interactive approaches. *Internet of Things*, *25*, 101048. https://doi.org/10.1016/j.iot.2023.101048

Hanafi, M., Katsis, Y., Jindal, I., & Popa, L. (2022). A comparative analysis between human-in-the-loop systems and large language models for pattern extraction tasks. In *Proceedings of the Fourth Workshop on Data Science with Human-in-the-Loop (Language Advances)* (pp. 43–50). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.dash-1.7

Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, *3*(2), 119–131. https://doi.org/10.1007/s40708-016-0042-6

Lazaros, K., Vrahatis, A. G., & Kotsiantis, S. (2026). Human-in-the-loop artificial intelligence: A systematic review of concepts, methods, and applications. *Entropy, 28*(4), 377. https://doi.org/10.3390/e28040377

Liu, H., Nasiriany, S., Zhang, L., Bao, Z., & Zhu, Y. (2025). Robot learning on the job: Human-in-the-loop autonomy and learning during deployment. *The International Journal of Robotics Research*, *44*(10–11), 1727–1742. https://doi.org/10.1177/02783649241273901

Liu, H. K., Tang, M., & Collard, A. S. J. (2025). Hybrid intelligence for the public sector: A bibliometric analysis of artificial intelligence and crowd intelligence. *Government Information Quarterly*, *42*(1), 102006. https://doi.org/10.1016/j.giq.2024.102006

Memarian, B., & Doleck, T. (2024). Human-in-the-loop in artificial intelligence in education: A review and entity-relationship (ER) analysis. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100053. https://doi.org/10.1016/j.chbah.2024.100053

Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2023). Human-in-the-loop machine learning: A state of the art. *Artificial Intelligence Review*, *56*(4), 3005–3054. https://doi.org/10.1007/s10462-022-10246-w

Olawade, D. B., Plabon, S. B., Ojo, A., Ogunbona, M. A., Makanjuola, B. D., & Olasilola, O. R. (2026). Human in the loop artificial intelligence in healthcare: Applications, outcomes, and implementation challenges. *International Journal of Medical Informatics*, *213*, 106362. https://doi.org/10.1016/j.ijmedinf.2026.106362

Retzlaff, C. O., Das, S., Wayllace, C., Mousavi, P., Afshari, M., Yang, T., Saranti, A., Angerschmid, A., Taylor, M. E., & Holzinger, A. (2024). Human-in-the-loop reinforcement learning: A survey and position on requirements, challenges, and opportunities. *Journal of Artificial Intelligence Research*, *79*, 359–415. https://doi.org/10.1613/jair.1.15348

Salloch, S., & Eriksen, A. (2024). What are humans doing in the loop? Co-reasoning and practical judgment when using machine learning-driven decision aids. *The American Journal of Bioethics*, *24*(9), 67–78. https://doi.org/10.1080/15265161.2024.2353800

Sele, D., & Chugunova, M. (2024). Putting a human in the loop: Increasing uptake, but decreasing accuracy of automated decision-making. *PLOS ONE*, *19*(2), e0298037. https://doi.org/10.1371/journal.pone.0298037

Toumanidis, L., Kasnesis, P., Chatzigeorgiou, C., Feidakis, M., & Patrikakis, C. (2021). ActiveCrowds: A human-in-the-loop machine learning framework. In C. Frasson, K. Kabassi, & A. Voulodimos (Eds.), *Novelties in intelligent digital systems: Proceedings of the 1st International Conference (NIDS 2021), Athens, Greece, September 30 – October 1, 2021* (pp. 176–184). IOS Press. https://doi.org/10.3233/FAIA210090

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems*, *135*, 364–381. https://doi.org/10.1016/j.future.2022.05.014