ReGeo ISSN: 2177-3246

GEOGRAPHIC CONCENTRATION, SECTORAL DISTRIBUTION, AND REPORTING TRANSPARENCY IN HUMAN-AI COLLABORATION RESEARCH: A CORPUS AUDIT OF 438 STUDIES (2018 TO 2026)

CONCENTRAÇÃO GEOGRÁFICA, DISTRIBUIÇÃO SETORIAL E TRANSPARÊNCIA DE RELATO NA PESQUISA EM COLABORAÇÃO HUMANO-IA: UMA AUDITORIA DE CORPUS DE 438 ESTUDOS (2018 TO 2026)

CONCENTRACIÓN GEOGRÁFICA, DISTRIBUCIÓN SECTORIAL Y TRANSPARENCIA DE REPORTE EN LA INVESTIGACIÓN EN COLABORACIÓN HUMANO-IA: UNA AUDITORÍA DE CORPUS DE 438 ESTUDIOS (2018 TO 2026)

https://doi.org/10.56238/revgeov17n5-158

Data de submissão: 16/04/2026 **Data de publicação:** 16/05/2026

Marcelo Caetano de Ribeiro e Melo[1], Marcos Alves da Rocha[2], Silvania Soares Brandão[3], Jordana Layara Alves Dias[4]

ABSTRACT

This paper audits the geographic and sectoral distribution of primary studies within a 438-study corpus on human-AI team collaboration, covering the period from 2018 to 2026. Two analytically distinct dimensions are examined: reporting transparency, defined as the rate at which studies disclose the national context and economic sector of data collection; and, restricted to studies that provide usable information, the degree of concentration across countries and sectors. Country of data collection was not reported or not applicable in 363 studies (82.9% of the corpus), severely limiting any inference about the geographic scope of current evidence. Among the 75 studies (17.1%) with usable country information, China and the United States together account for 42 cases (56.0% of usable observations), indicating marked concentration in two high-income settings. Sector information was absent in 134 studies (30.6%), whereas among the 304 studies (69.4%) with usable sector data, Business, Management and Finance concentrates 247 cases (81.3%). Healthcare, Education and Academia, and Experimental contexts each account for fewer than 5.4% individually. These findings raise substantive questions about the external validity of current human-AI collaboration findings and constitute a call for standardized, systematic reporting practices in the field. Methodological decisions regarding country normalization and sector categorization are fully documented to ensure reproducibility.

Keywords: Human-AI Collaboration. Geographic Concentration. Sectoral Distribution. Reporting Transparency. Corpus Audit.

[1] Ph.D. in Biofuel Science and Technology. Institution: Universidade Estadual de Montes Claros (UNIMONTES). E-mail: marcelo.melo@unimontes.br

[2] Master's degree in Economic Development and Business Strategy. Institution: Universidade Estadual de Montes Claros (UNIMONTES). E-mail: marcos.alves@unimontes.br

[3] Master's degree in Business Administration. Institution: Instituto Federal do Norte de Minas Gerais (IFNMG). E-mail: silvania.brandao@ifnmg.edu.br

[4] Specialist in Accounting Management and Corporate Controllership. Institution: Faculdades Integradas do Norte de Minas (FUNORTE). E-mail: jordana.dias@soufunorte.com.br

RESUMO
Este artigo audita a distribuição geográfica e setorial de estudos primários dentro de um corpus de 438 estudos sobre colaboração humano-IA, cobrindo o período de 2018 a 2026. Duas dimensões analiticamente distintas são examinadas: transparência de relatórios, definida como a taxa em que os estudos divulgam o contexto nacional e o setor econômico da coleta de dados; e, restrita aos estudos que fornecem informações utilizáveis, o grau de concentração entre países e setores. O país de coleta de dados não foi relatado ou não se aplicava em 363 estudos (82,9% do corpus), limitando severamente qualquer inferência sobre o escopo geográfico das evidências atuais. Entre os 75 estudos (17,1%) com informações de país utilizáveis, China e Estados Unidos juntos representam 42 casos (56,0% das observações utilizáveis), indicando concentração marcada em dois contextos de alta renda. As informações de setor estavam ausentes em 134 estudos (30,6%), enquanto entre os 304 estudos (69,4%) com dados de setor utilizáveis, Negócios, Gestão e Finanças concentra 247 casos (81,3%). Saúde, Educação e Academia, e contextos Experimentais respondem cada um por menos de 5,4% individualmente. Esses achados levantam questões substantivas sobre a validade externa das descobertas atuais de colaboração humano-IA e constituem um apelo por práticas de relatórios padronizadas e sistemáticas no campo. As decisões metodológicas com relação à normalização de países e categorização de setores estão totalmente documentadas para garantir a reprodutibilidade.

Palavras-chave: Colaboração Humano-IA. Concentração Geográfica. Distribuição Setorial. Transparência de Relatórios. Auditoria de Corpus.

RESUMEN
Este artículo audita la distribución geográfica y sectorial de estudios primarios dentro de un corpus de 438 estudios sobre colaboración humano-IA, cubriendo el período de 2018 a 2026. Se examinan dos dimensiones analíticamente distintas: transparencia de informes, definida como la tasa en que los estudios divulgan el contexto nacional y el sector económico de la recopilación de datos; y, restringida a estudios que proporcionan información utilizable, el grado de concentración entre países y sectores. El país de recopilación de datos no fue reportado o no era aplicable en 363 estudios (82,9% del corpus), limitando severamente cualquier inferencia sobre el alcance geográfico de la evidencia actual. Entre los 75 estudios (17,1%) con información de país utilizable, China y Estados Unidos juntos representan 42 casos (56,0% de observaciones utilizables), indicando una concentración marcada en dos contextos de altos ingresos. La información del sector estuvo ausente en 134 estudios (30,6%), mientras que entre los 304 estudios (69,4%) con datos de sector utilizables, Negocios, Gestión y Finanzas concentra 247 casos (81,3%). Salud, Educación y Academia, y contextos Experimentales representan cada uno menos del 5,4% individualmente. Estos hallazgos plantean preguntas sustanciales sobre la validez externa de los hallazgos actuales de colaboración humano-IA y constituyen un llamado a prácticas de informes estandarizadas y sistemáticas en el campo. Las decisiones metodológicas con respecto a la normalización de países y categorización de sectores están completamente documentadas para garantizar la reproducibilidad.

Palabras clave: Colaboración Humano-IA. Concentración Geográfica. Distribución Sectorial. Transparencia de Informes. Auditoría de Corpus.

1 INTRODUCTION

The emergence of artificial intelligence as an active participant in organizational work processes has generated substantial scholarly attention over the past decade. From autonomous decision-support agents to large language models deployed as cognitive collaborators, AI systems have moved beyond tool status toward configurations that researchers increasingly characterize as hybrid teams, joint cognitive systems, or collaborative partnerships between human and machine intelligence (DELLERMANN *et al.*, 2018; HAGEMANN *et al.*, 2023; FABRI *et al.*, 2023). The intellectual momentum behind this shift is reflected in a rapidly growing empirical literature: studies on human-AI collaboration now span computer science, organizational behavior, human-computer interaction, information systems, healthcare informatics, educational technology, and manufacturing engineering, among other disciplinary homes.

Yet rapid growth in a research field does not automatically translate into a well-distributed or representative evidence base. Research communities tend to cluster around established institutions, funding streams, publication traditions, and economic contexts. When primary studies concentrate in a narrow range of countries or sectors, the theoretical constructs and design principles that emerge from that literature reflect, implicitly or explicitly, the particular social, organizational, and institutional conditions of those settings. Findings derived overwhelmingly from large-scale corporate environments in two or three high-income economies may not transfer readily to small-enterprise contexts, to public-sector organizations in emerging economies, or to sectors where human-AI interaction takes markedly different forms.

A second and equally important concern is reporting transparency. Before any inference about geographic or sectoral concentration can be drawn, it is necessary to establish what fraction of the available evidence base actually discloses the information needed to make such inferences. If the country in which data collection took place is not reported in the majority of studies, concentration estimates derived from the reporting minority may be substantially misleading. Transparency in reporting the provenance and context of empirical data is, therefore, a prerequisite for cumulative scientific progress, not merely a methodological nicety.

Against this backdrop, the present study pursues two interconnected objectives. The first is descriptive and diagnostic: to determine the rate at which studies within a 438-study corpus on human-AI team collaboration disclose the national context and economic sector of

their empirical work, treating non-disclosure as a substantive finding in its own right. The second objective is distributional: among the subset of studies that do provide usable geographic and sectoral information, to characterize the degree of concentration and to identify which countries and sectors are represented, underrepresented, or absent entirely.

These objectives are addressed through a systematic corpus audit. The analytical corpus comprises 438 primary studies retrieved through structured bibliographic searches covering the period from 2018 to 2026. One study present in the original source database was excluded because it involved no real AI system, leaving a final analytical sample of 438 unique studies. No new database search was conducted for this work; the corpus is entirely inherited from those prior systematic searches, and this study functions as a methodological audit of the evidence base those searches produced.

The paper makes three substantive contributions. First, it documents a high rate of country non-reporting (82.2% of the corpus), which constitutes a transparency deficit that should be addressed by reporting standards in the field. Second, it demonstrates that, among the minority of studies with usable country information, geographic concentration is substantial, with China and the United States collectively accounting for more than half of observations. Third, it reveals extreme sectoral concentration: more than four in five studies with usable sector data originate from Business, Management and Finance contexts, leaving sectors such as manufacturing, creative industries, and military-civilian operations virtually unexamined in this literature.

The remainder of this paper is organized as follows. Section 2 reviews the theoretical background, covering the evolution of human-AI teaming concepts, what is known about geographic and sectoral biases in empirical research, and the role of reporting transparency. Section 3 describes the methodology, including the source corpus, the country normalization procedure, the sector categorization scheme, and the analytical approach. Section 4 presents and discusses results for reporting transparency, geographic distribution, and sectoral distribution. Section 5 draws conclusions, acknowledges limitations, and identifies directions for future inquiry. Country normalization rules are documented in full in Appendix A, and the sector categorization scheme appears in Appendix B.

2 THEORETICAL BACKGROUND

2.1 THE CONCEPTUAL EVOLUTION OF HUMAN-AI TEAMING

The concept of human-AI collaboration encompasses a wide range of configurations in which artificial intelligence systems contribute cognitive, communicative, or physical capabilities alongside human actors toward shared goals. Early theoretical treatments tended to position AI as a tool or decision-support instrument, subordinate to human agency and evaluated primarily by task performance criteria (ROSENFELD; RICHARDSON, 2019; LEVINE; WILLIAMS, 2018). Over time, as AI systems acquired greater autonomy, adaptability, and communicative competence, scholars began to develop richer conceptual frameworks that recognized AI as a functional team member rather than merely a sophisticated instrument.

DELLERMANN *et al.* (2018) proposed the construct of hybrid intelligence, defined as the ability of a joint human-machine system to achieve goals beyond the capability of either component acting alone. This formulation places human-AI collaboration within a broader tradition of research on collective and distributed cognition, while also anticipating the organizational design challenges that arise when artificial agents participate alongside humans in consequential decisions. The hybrid intelligence framework was subsequently extended and refined by several research streams. FABRI *et al.* (2023) disentangled the multiple ways in which human and AI contributions can be combined, distinguishing configurations by degree of integration, temporal ordering, and locus of authority. SIEMON (2022) elaborated a taxonomy of AI team roles, arguing that the specific function an AI system performs within a team, ranging from coordinator to advisor to executor, shapes team dynamics in ways that existing organizational role theory had not anticipated. HAGEMANN *et al.* (2023) foregrounded the challenges that team-centered AI creates for workers, supervisors, and organizations, highlighting the need for team-level rather than individual-level models of human-AI interaction. LI and TIAN (2026) synthesized evidence on AI-human collaboration in decision-making, proposing a conceptual framework that integrates task characteristics, human factors, and AI capabilities as jointly determinative of collaboration outcomes.

The body of empirical work that has accumulated around these frameworks is substantial. GEORGANTA, PEUS and NIESS (2023) conducted a systematic review of interactive technologies through the lens of team effectiveness, identifying trust, communication, and role clarity as recurring mediators and moderators. PUERTA-

BELDARRAIN *et al.* (2025) provided a comprehensive review of the human-AI collaboration landscape, mapping theoretical perspectives, empirical methods, and key findings across multiple disciplines. DELL'ANNA *et al.* (2026) examined quality attributes that distinguish high-performing hybrid intelligence teams from less effective ones, while FÜGENER *et al.* (2021) investigated the risk of humans becoming excessively dependent on AI judgment, a dynamic they termed “Borg” integration.

Taken together, this body of work has established human-AI teaming as a coherent and productive research area. However, the very productivity of the field invites the question of whether its empirical foundations are sufficiently broad. Theoretical frameworks developed from studies conducted in a limited number of countries or economic sectors may carry implicit cultural, institutional, or organizational assumptions that reduce their applicability to other contexts.

2.2 GEOGRAPHIC CONCENTRATION AND THE GENERALIZABILITY CHALLENGE

The problem of geographic concentration in empirical social and behavioral science has been discussed extensively since HENRICH, HEINE and NORENZAYAN (2010) demonstrated that participants in a disproportionate share of behavioral research originate from Western, Educated, Industrialized, Rich, and Democratic societies. The WEIRD critique, as it became known, challenged the disciplinary assumption that findings from convenience samples drawn predominantly from North American and European university populations could serve as a universal baseline for human psychology. Subsequent scholarship extended the critique to other domains, including organizational research, technology studies, and human-computer interaction, where geographic concentration raises parallel questions about the degree to which findings reflect universal regularities versus contextually specific patterns.

In the human-AI collaboration literature, geographic concentration takes a specific form. AI systems are developed, deployed, and used unevenly across national and regional contexts. The institutional environments, regulatory frameworks, cultural orientations toward technology, and economic conditions that shape human-AI interaction differ substantially between, for example, a large technology firm in coastal China, a German automotive manufacturing plant, a Jordanian healthcare institution, and a Colombian creative agency. If empirical research on human-AI teaming is concentrated in a small number of national

settings, the findings it produces may systematically under-represent the institutional and cultural heterogeneity that characterizes AI deployment in practice.

The issue is compounded by the nature of human-AI collaboration research, which typically requires access to organizations actively deploying AI systems. Such organizations are not uniformly distributed across the world. Countries at the frontier of AI development, particularly China and the United States, host a disproportionate share of AI-adopting enterprises and may therefore attract a disproportionate share of empirical research simply because access is easier. This creates a concentration mechanism that operates independently of any explicit geographic selection bias.

OFOSU-AMPONG and RENNER-MICAH (2026) addressed this issue directly, examining the cross-cultural acceptability of AI as a co-worker across multiple national and cultural settings. Their findings underscore the degree to which human-AI collaboration is culturally embedded: attitudes toward AI authority, trust calibration, and collaborative norms varied substantially across contexts. This suggests that theories and design principles derived from a narrow geographic base may carry hidden cultural contingencies that only become visible when research is extended to more diverse settings.

Reporting transparency is a prerequisite for diagnosing geographic concentration. If the country of data collection is not disclosed in a substantial fraction of studies, any concentration estimate based on disclosed information is potentially unrepresentative of the full corpus. This paper treats the non-reporting rate as a primary empirical finding, not merely a methodological inconvenience to be resolved before proceeding to distributional analysis.

2.3 SECTORAL DISTRIBUTION AND THE SCOPE OF EVIDENCE

Sectoral distribution in empirical research is governed by a combination of access, funding, and disciplinary affinity. Organizational behavior and information systems research, two of the largest contributors to human-AI collaboration scholarship, have traditionally centered on business and management contexts: firms, workplaces, managerial decision-making, and organizational performance. This disciplinary anchoring tends to produce sectoral concentration in business and management settings even when the research questions being addressed are, in principle, applicable to a much broader range of sectors.

At the same time, several sectors have attracted dedicated human-AI collaboration research precisely because the stakes of human-AI interaction are particularly high. Healthcare presents one of the most extensively studied specialized contexts. HENRY *et al.*

(2022) documented clinician experiences with a deployed machine learning system, finding that human-machine teaming is critical to AI adoption in clinical settings and that trust, transparency, and workload are key determinants of effective collaboration. BOSSEN and PINE (2023) examined human-AI collaboration in clinical documentation work, finding that healthcare professionals develop sophisticated division-of-labor strategies with AI systems that differ substantially from the patterns observed in business decision-support contexts. EVERETT *et al.* (2026) conducted a randomized controlled trial of clinician-AI collaborative diagnostic workflows, producing evidence on the conditions under which AI assistance improves diagnostic accuracy without generating automation bias.

Manufacturing and industrial production represent a second specialized context where human-AI teaming has attracted substantive attention. BOCKLISCH *et al.* (2024) examined hybrid decision-making in atmospheric plasma spraying, a context in which AI systems support real-time quality control alongside human operators. KRAUSE *et al.* (2024) addressed the organizational challenges of managing human-AI collaborations within Industry 5.0 scenarios, where AI systems are expected to support human workers rather than replace them. KRZYWDZINSKI *et al.* (2025) conducted an experimental study on human-AI decision-making in the context of automation failures, finding that team organization significantly affects the ability to handle unexpected AI errors, a finding with immediate implications for manufacturing safety.

Education and creative industries represent additional specialized contexts. In educational settings, AI systems are increasingly deployed as tutoring agents, learning companions, and team collaborators in group learning tasks (MARRONE *et al.*, 2024; HU; CHAN, 2025). In creative industries, AI systems contribute as co-authors, co-designers, and creative partners, raising questions about attribution, authorship, and the nature of creativity itself (REZWANA; MAHER, 2023; JIANG *et al.*, 2024).

Despite these specialized research streams, the bulk of human-AI collaboration research appears concentrated in business and management contexts. The degree of this concentration, and the extent to which it limits the scope of current evidence, is a central empirical question addressed in the present study.

2.4 REPORTING TRANSPARENCY AS A METHODOLOGICAL STANDARD

Reporting transparency in empirical research encompasses the disclosure of information that allows readers to assess the scope, context, and generalizability of study

findings. For empirical work on human-AI collaboration, relevant transparency attributes include the country or region in which data were collected, the economic sector or organizational context from which participants were drawn, the AI system(s) studied, the temporal context of data collection, and the characteristics of human participants. These disclosures enable both replication and cumulative synthesis across studies.

The PRISMA 2020 guidelines (PAGE *et al.*, 2021), developed for systematic reviews, emphasize reporting completeness as a foundational requirement for scientific transparency. Although PRISMA addresses systematic reviews rather than primary studies, its underlying logic applies to any empirical research tradition that aspires to cumulative knowledge-building: findings that cannot be contextualized cannot be integrated. When country and sector information are absent from primary studies, synthesis across studies must proceed without knowledge of what contexts are actually covered by the evidence base.

In the human-AI collaboration literature, country and sector non-reporting may arise for several reasons. Authors may consider geographic context irrelevant or self-evident (as when a study is conducted at a named institution whose nationality is apparent from the affiliation). They may not distinguish between the location of the research team and the location of data collection. Organizational confidentiality constraints may prevent disclosure of the sector in which a corporate partner operates. In laboratory-based experimental research, the concept of a “sector” may appear inapplicable, since participants may be drawn from a general student population without representing any specific economic sector.

These reasons explain non-reporting without justifying it. From the perspective of cumulative evidence synthesis, the absence of country and sector information in a study represents a transparency deficit that reduces the value of that study's contribution to the broader literature. Documenting the prevalence of this deficit across the human-AI collaboration corpus is, therefore, a scientifically and practically significant undertaking.

3 METHODOLOGY

3.1 SOURCE CORPUS

The analytical corpus used in this study comprises 438 unique primary studies on human-AI team collaboration. The corpus was assembled through structured bibliographic searches conducted across multiple academic databases, covering the period from 2018 to 2026. The year 2018 marks an approximate threshold at which the human-AI teaming literature began to consolidate as a distinct research stream, coinciding with the emergence

of foundational theoretical frameworks that explicitly conceptualized AI as a team member rather than a tool (DELLERMANN *et al.*, 2018; LEVINE; WILLIAMS, 2018). The year 2026 represents the upper boundary of the search window.

The source database contained records for 439 unique studies after deduplication. One study was excluded because it involved no real AI system, meaning the described collaborative configuration did not include an operational artificial intelligence component. The final analytical corpus therefore comprises 438 unique primary studies.

For each study, the following fields were extracted and retained: full bibliographic identifier, title, year of publication, country of data collection (as reported by the authors), and the economic sector or organizational context from which empirical data were drawn. The analysis presented here operates exclusively on the country and sector fields, treating all other bibliographic information as context.

It is important to distinguish the present study from a new independent database search. No new literature search was conducted. The 438-study corpus is inherited in its entirety from prior systematic searches. The contribution of this paper consists of the systematic audit and normalization of the geographic and sectoral metadata within that corpus, and the analysis of the distributional patterns those normalized data reveal.

3.2 COUNTRY NORMALIZATION PROCEDURE

The raw country field in the source database contains entries of highly variable form and informational content. Entries range from standardized country names in English or Portuguese to ambiguous descriptive phrases, multi-country lists, and explicit statements of non-disclosure or non-applicability. A multi-step normalization procedure was applied to convert these heterogeneous entries into a consistent categorical variable.

Step 1: Identification of explicit non-reporting. Entries containing variants of “not reported,” “not specified,” “not identified,” or equivalent phrases in Portuguese and English were coded as *Not reported*. Entries that were left blank or that contained only generic filler text were similarly coded as *Not reported*. Ambiguous entries that contained contextual clues but did not confirm a specific country of data collection (for example, entries noting that the use of a US regulatory framework such as HIPAA suggests a US context, without confirming that data collection actually took place there) were coded as *Not reported* rather than attributed to a specific country.

Step 2: Identification of not-applicable cases. Entries referring to international corpora, multi-country indexing databases, or corpus-level indexing resources (for example, “not applicable; international corpus indexed in Web of Science”) were coded as *Not applicable*. This category reflects studies of a theoretical or review character for which a single country of data collection is genuinely absent, not simply undisclosed.

Step 3: Identification of multi-country entries. Entries containing comma-separated lists of two or more country names, explicit Portuguese-language phrases indicating multiple countries (“múltiplos países,” “vários países”), or large-scale multi-national expressions (“80 países”) were coded as *Multiple countries*. This category captures studies that intentionally drew participants from multiple national contexts and should not be conflated with studies for which the country is simply unknown.

Step 4: Language-specific and institutional references. Entries referring to participant attributes that imply a national context (for example, “Chinese professionals” in Portuguese: “profissionais chineses”) were resolved to the corresponding country. Institutional abbreviations that unambiguously identify a national context (for example, CSIRO, the Australian national science agency) were, however, treated with caution: when the institutional reference appeared in a field that had otherwise been flagged as unspecified, the entry was retained as *Not reported* on the grounds that the institutional location of the research team does not necessarily correspond to the country of data collection.

Step 5: Standardization of confirmed country names. Country names in Portuguese, English, or abbreviated form were mapped to standardized English country names (for example, “Alemanha” to “Germany,” “Estados Unidos” to “United States,” “EUA” to “United States”).

The resulting normalized variable takes the value *Not reported*, *Not applicable*, *Multiple countries*, or a specific country name. For the purposes of the reporting transparency analysis (Section 4.1), *Not reported* and *Not applicable* are jointly classified as “not usable,” following the logic that neither provides actionable information about geographic scope. For the concentration analysis (Section 4.2), only the 75 studies coded to a specific country or to *Multiple countries* are included. The full mapping of original raw entries to normalized categories is provided in Appendix A.

3.3 SECTOR CATEGORIZATION SCHEME

The raw sector field in the source database presents challenges analogous to those in the country field: entries combine formal sector designations, informal descriptions of organizational context, study design characteristics, and explicit statements of non-disclosure. An eight-category scheme was developed to accommodate the range of entries observed, structured with an explicit priority ordering to ensure consistent assignment of ambiguous cases.

The eight substantive categories are as follows:

1. Healthcare and Medicine: studies in which the primary empirical setting involves clinical care, medical diagnosis, health services management, or health information systems.
2. Military, Defense and Public Safety: studies whose empirical context involves defense organizations, law enforcement, emergency response, or combat simulation.
3. Manufacturing and Industrial Production: studies set in production, assembly, quality control, logistics, or industrial engineering environments.
4. Technology and Software Development: studies whose primary context is software engineering, systems development, or technology product design, where the human-AI collaboration operates within a technology creation process rather than as a management support function.
5. Education and Academia: studies conducted in educational institutions, involving students, instructors, or educational processes as the primary empirical focus.
6. Creative Industries: studies in design, music, visual arts, fashion, journalism, or other creative domains where the human-AI collaboration centers on creative production.
7. Experimental or Generic Task: laboratory or controlled experimental studies in which participants perform abstract or generic tasks not tied to a specific real-world sector, and studies that use experimental paradigms drawn from cognitive psychology or behavioral science rather than a particular organizational sector.
8. Business, Management and Finance: a broad residual category capturing studies set in corporate, managerial, financial services, consulting, or generic organizational contexts, as well as studies whose organizational context is described only in terms of managerial or decision-making activities without a more specific sectoral designation.

The priority ordering matters for ambiguous cases. Healthcare and Military are assessed before Manufacturing, to ensure that studies in medical device manufacturing or defense logistics are assigned to the more specific category. Creative is assessed before Experimental, since creative lab studies should be categorized by their domain rather than their method. Business, Management and Finance is assessed last and functions as the broadest fallback category, capturing the large volume of studies that describe organizational or managerial contexts without a more specific sectoral anchor.

Entries that contained no usable sectoral information, including blank fields and entries equivalent to “not reported” or “not specified,” were coded as *Not reported*. The full sector categorization mapping is provided in Appendix B.

3.4 ANALYTICAL APPROACH

The analysis is organized in two stages. The first stage examines reporting transparency: the share of the 438-study corpus for which country and sector information is usable versus not usable. This stage treats the full corpus as the denominator and reports counts and percentages for each category. The transparency analysis is presented as a primary finding in its own right, not merely as a data quality caveat preceding the distributional analysis.

The second stage examines concentration: the distribution of studies across countries and sectors, restricted to the subset of studies for which usable information exists. For this stage, the relevant denominators are 75 (studies with usable country information) and 304 (studies with usable sector information), not 438. All percentages in the concentration analysis are computed relative to the usable subset, not the full corpus. This distinction is maintained throughout the results section to prevent misrepresentation of what can and cannot be inferred from available data.

No inferential statistical tests are applied. The study is descriptive in orientation: its goal is to characterize the distribution of available evidence, not to test hypotheses about the mechanisms driving that distribution. Concentration is described through counts and percentages, supplemented by ranked displays. Two figures summarize the distributional findings visually. All normalization and analysis scripts were written in Python and are fully reproducible from the source data.

4 RESULTS AND DISCUSSION

4.1 REPORTING TRANSPARENCY: COUNTRY AND SECTOR NON-DISCLOSURE RATES

Table 1 and Table 2 present the full distributions for country and sector, respectively, across all 438 studies in the corpus. The findings regarding reporting transparency are striking.

Table 1

Geographic reporting status across the full corpus (N = 438)

Geographic Category	n	% of corpus
Not reported or not applicable	363	82.9
China	29	6.6
United States	13	3.0
Multiple countries	10	2.3
Germany	4	0.9
Japan	2	0.5
Australia	2	0.5
Netherlands	2	0.5
United Kingdom	2	0.5
India	2	0.5
Jordan	1	0.2
Ecuador	1	0.2
Ireland	1	0.2
Denmark	1	0.2
Finland	1	0.2
Kosovo	1	0.2
Colombia	1	0.2
Kuwait	1	0.2
Taiwan	1	0.2
South Korea	1	0.2
Total	**438**	**100.0**

"Not applicable" denotes theoretical or review studies for which a single country of data collection does not exist. "Not reported" denotes empirical studies that did not disclose the country of data collection. Both categories are jointly classified as not usable for concentration analysis.
Source: Prepared by the authors.

Country of data collection is not usable in 363 studies, representing 82.9% of the corpus (360 not reported plus 3 not applicable). Only 75 studies, or 17.1% of the corpus, provide usable country information. This non-reporting rate is the first and arguably the most consequential finding of this analysis. Whatever geographic concentration patterns may exist in the human-AI collaboration literature, they can be directly observed only for a small minority of the evidence base.

Table 2

Sectoral reporting status across the full corpus (N = 438)

Sector Category	n	% of corpus
Not reported or not applicable	134	30.6
Business, Management and Finance	247	56.4
Healthcare and Medicine	16	3.7
Education and Academia	16	3.7
Experimental or Generic Task	15	3.4
Manufacturing and Industrial Production	4	0.9
Technology and Software Development	4	0.9
Creative Industries	1	0.2
Military, Defense and Public Safety	1	0.2
Total	**438**	**100.0**

Source: Prepared by the authors.

Sector information is not reported in 134 studies, or 30.6% of the corpus. The sector non-reporting rate is substantially lower than the country non-reporting rate, which suggests that study authors are more likely to disclose the organizational context from which their empirical data were drawn than the national context. This differential reporting rate may reflect the fact that sector is often implied by the venue, task, or participant description (for example, a study on AI-assisted financial forecasting implicitly discloses a business and finance context even without an explicit statement), whereas country may be considered irrelevant or sensitive by some authors.

The contrast between the country non-reporting rate (82.9%) and the sector non-reporting rate (30.6%) is analytically important. It means that the sector-level evidence base is more complete than the country-level evidence base, but it also means that nearly one

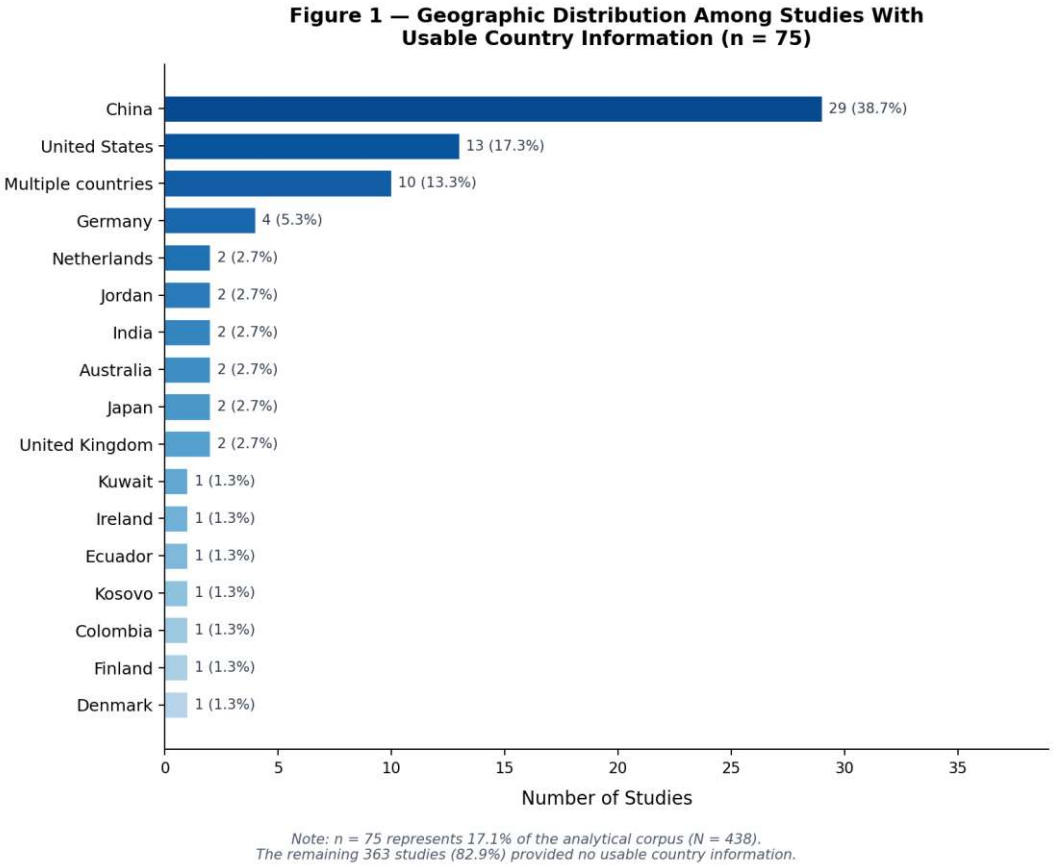
third of the corpus provides no information about the sectoral context of human-AI collaboration. Future reporting standards should address both dimensions.

4.2 GEOGRAPHIC DISTRIBUTION AMONG STUDIES WITH USABLE COUNTRY INFORMATION

Among the 75 studies with usable country information, the distribution is highly concentrated. Figure 1 illustrates the ranked distribution across 17 geographic categories (16 individual countries plus the multiple-countries category).

Figure 1

Geographic distribution of studies with usable country information (n = 75 usable observations)



Source: generated from analytical corpus, 2026. Full distribution given in Table 1.

China leads with 29 studies (38.7% of the usable subset), followed by the United States with 13 studies (17.3%). Together, these two countries account for 42 of the 75 usable observations, or 56.0% of the geographically identified evidence base. The dominance of China and the United States is consistent with their position as the two largest investors in AI

development globally, but it means that more than half of the geographically attributable human-AI collaboration evidence derives from two specific national settings, each with distinctive institutional, regulatory, and cultural characteristics that are not necessarily shared by the broader global landscape of AI deployment.

The multi-country category adds 10 studies (13.3%), bringing the combined share of China, United States, and multi-country studies to 52 cases (69.3% of the usable subset). These multi-country studies are valuable because they explicitly draw participants from diverse national contexts; however, they also tend to aggregate rather than differentiate national patterns, limiting their utility for country-specific inference.

Germany is the next most frequently represented single country, with 4 studies (5.3%), followed by six countries with 2 studies each: Japan, Australia, Netherlands, United Kingdom, India, and Jordan. Seven additional countries appear exactly once each: Ecuador, Ireland, Denmark, Finland, Kosovo, Colombia, and Kuwait. Collectively, all non-China, non-US single-country studies contribute 33 cases (44.0% of the usable subset), and within this group the representation of the Global South is limited to Jordan (2), Ecuador (1), Colombia (1), and Kuwait (1), a total of 5 cases from outside the high-income Western and East Asian cluster.

These patterns carry implications that extend beyond the human-AI collaboration literature into broader debates about the generalizability of organizational and behavioral research. The WEIRD critique documented by HENRICH, HEINE and NORENZAYAN (2010) applies here with particular force: studies conducted in high-income, technologically advanced, and institutionally stable settings may capture human-AI collaboration dynamics that are specific to those settings rather than universal. The relationship between human workers and AI systems is shaped, among other things, by employment law, organizational culture, workforce education levels, attitudes toward automation, and trust in technological institutions, all of which vary substantially across the countries represented in this corpus and, even more dramatically, across the countries that are absent.

The near-total absence of sub-Saharan African countries, South and Southeast Asian countries (India appears, but with only 2 studies), Central Asian, and most Latin American countries from the usable portion of the evidence base is particularly notable. These regions collectively contain the majority of the world's workforce and are sites of rapidly expanding AI deployment in sectors ranging from agriculture to financial services. Research that does not engage empirically with these contexts cannot speak to the human-AI collaboration challenges and opportunities they present.

Even within the regions that are represented, diversity is limited. The European contingent consists of Germany (4), Netherlands (2), United Kingdom (2), Ireland (1), Denmark (1), Finland (1), and Kosovo (1), a total of 12 studies drawn predominantly from Northwestern Europe. Southern and Eastern Europe are largely absent. Within Asia, China (29) and Japan (2) are the represented countries aside from India (2) and Jordan (2), with Southeast Asia, Central Asia, and the Korean peninsula appearing only within multi-country studies if at all.

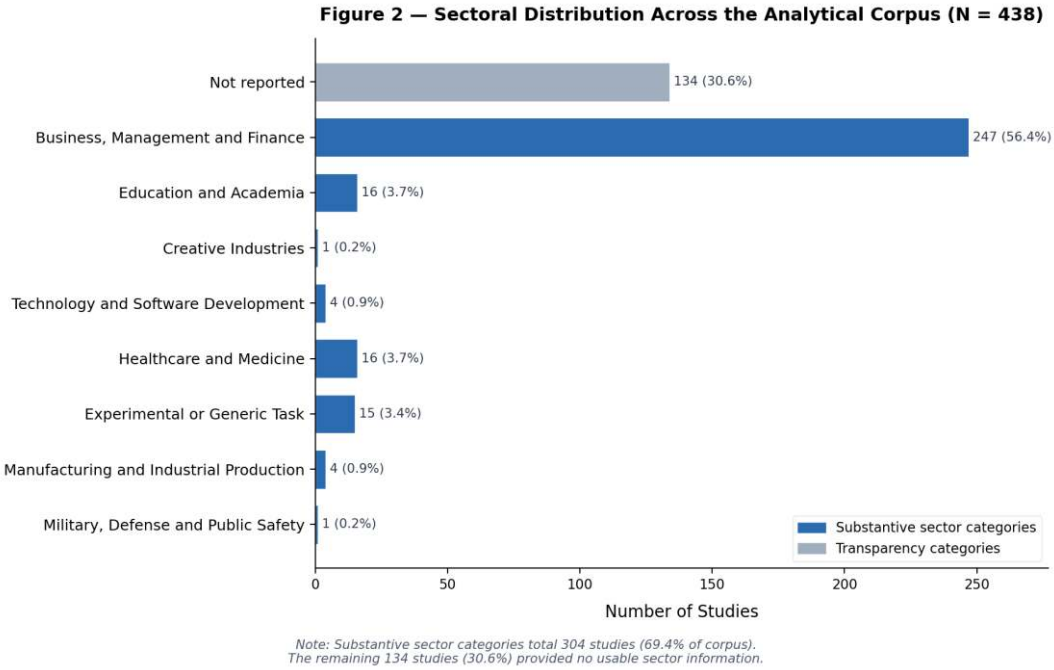
The temporal distribution of these 75 studies across the 2018 to 2026 window was not systematically analyzed in this study, but the concentration of recent publications from China and the United States (a substantial share of studies with 2025 and 2026 publication dates originate from these two countries, as visible in the source data) suggests that geographic concentration is not diminishing as the field matures. If anything, the acceleration of AI research capacity in China and the United States may be reinforcing rather than reducing the concentration trend.

4.3 SECTORAL DISTRIBUTION AMONG STUDIES WITH USABLE SECTOR INFORMATION

Among the 304 studies with usable sector information, the sectoral distribution is even more sharply concentrated than the geographic distribution. Figure 2 illustrates the ranked sectoral distribution.

Figure 2

Sectoral distribution of studies with usable sector information (n = 304 usable observations)



Source: generated from analytical corpus, 2026. Full distribution given in Table 2.

Business, Management and Finance dominates with 247 studies, representing 81.3% of the usable sector subset. This concentration is remarkable: four in five studies with disclosed sectoral context originate from a single broad sector. Every other sector accounts for fewer than 5.4% of usable observations individually.

Education and Academia and Healthcare and Medicine are tied at 16 studies each (5.3% each). Experimental or Generic Task settings contribute 15 studies (4.9%). Manufacturing and Industrial Production and Technology and Software Development account for 4 studies each (1.3%). Creative Industries and Military, Defense and Public Safety are represented by a single study each (0.3%).

The dominance of Business, Management and Finance must be interpreted with appropriate nuance. The sector is genuinely broad: it encompasses studies conducted in financial services firms, consulting practices, corporate strategy contexts, supply chain management, human resources, marketing, product development, and generic organizational decision-making. The sector's apparent dominance partly reflects its definitional scope. A narrower definition might redistribute some studies to more specific categories (a study on AI in financial forecasting might be coded as finance rather than

business management, for instance). However, even under a narrower definition, the concentration in organizational and managerial contexts would remain substantial.

Healthcare and Medicine is the most visible specialized sector at 5.3%, reflecting the dedicated research streams on clinical AI that have grown substantially in recent years (HENRY *et al.*, 2022; BOSSEN; PINE, 2023; EVERETT *et al.*, 2026). The healthcare human-AI collaboration literature is distinguished by its emphasis on real-world deployment studies, its attention to safety and accountability, and its use of randomized controlled trial methodologies that are uncommon elsewhere in the human-AI collaboration corpus. The 16 healthcare studies in this corpus are not merely a proportional slice of a larger literature but represent a methodologically and substantively distinct strand.

Education and Academia (16 studies, 5.3%) reflects the growing deployment of AI tutoring systems, learning companions, and AI-assisted group work tools in educational settings (MARRONE *et al.*, 2024; HU; CHAN, 2025). The educational human-AI collaboration literature faces distinctive challenges around participant age, institutional context, and the tension between AI assistance and pedagogical goals.

The Experimental or Generic Task category (15 studies, 4.9%) represents research conducted in laboratory or controlled settings using abstract tasks rather than real-world organizational contexts. Such studies are valuable for isolating causal mechanisms, but their sector-agnostic nature limits direct applicability to any specific deployment context.

The extremely low representation of Manufacturing and Industrial Production (4 studies, 1.3%) is surprising given the prominence of human-AI collaboration in the Industry 4.0 and Industry 5.0 discourse (BOCKLISCH *et al.*, 2024; KRAUSE *et al.*, 2024; KRZYWDZINSKI *et al.*, 2025). Manufacturing AI applications, which involve human workers collaborating with robots, autonomous quality-control systems, and AI-assisted production planning tools, present distinctive collaboration challenges related to physical safety, real-time decision demands, and the embodied nature of manufacturing work. The small number of empirical studies in this sector suggests that the academic research community has not yet engaged systematically with manufacturing human-AI collaboration, despite its practical significance.

Technology and Software Development (4 studies, 1.3%) is similarly underrepresented. Given that AI-assisted programming, code review, and software design are among the most widely deployed human-AI collaboration applications in contemporary

practice, the paucity of empirical studies in this sector represents a notable gap in the literature.

Creative Industries (1 study, 0.3%) and Military, Defense and Public Safety (1 study, 0.3%) are essentially absent from the usable sector evidence base. The creative industries finding is particularly notable given the recent proliferation of generative AI tools in design, music, journalism, and the visual arts, contexts that have attracted considerable public attention and policy concern. The single military study in the corpus is consistent with the inherent difficulties of conducting and publishing empirical research on human-AI collaboration in defense settings, where access constraints and classification requirements limit scholarly inquiry.

4.4 DISCUSSION: IMPLICATIONS OF GEOGRAPHIC AND SECTORAL CONCENTRATION

The results presented in sections 4.1 through 4.3 converge on a coherent picture of the human-AI collaboration evidence base as it stood at the time of this analysis. That picture is characterized by two cross-cutting features: high rates of contextual non-reporting (82.9% for country, 30.6% for sector) and, among studies that do report contextual information, pronounced concentration in a small number of countries and a single broad sector.

The analytical and practical implications of this situation are substantial and deserve extended consideration.

Generalizability and the scope of evidence. The human-AI collaboration literature has produced a rich body of findings on trust dynamics, role allocation, team performance, and organizational outcomes. However, the geographic and sectoral evidence base from which these findings are drawn is narrow. Conclusions about how humans and AI systems interact effectively in teams are predominantly drawn from business and management settings in China and the United States, with secondary contributions from Europe and a small number of other countries. Generalizing these findings to manufacturing workers in Brazil, healthcare professionals in Indonesia, creative professionals in Egypt, or military operators in Finland requires assumptions about cross-contextual transferability that the current evidence base cannot validate.

This is not merely an academic concern about statistical generalization. It has direct consequences for the design of AI systems intended for deployment in diverse real-world settings. AI systems designed and evaluated for business and management contexts in high-

income countries may perform differently, and may interact with human collaborators differently, when deployed in other sectors and regions. Design principles for human-AI teaming that work well in a Chinese technology firm or an American consulting practice may need substantial revision before they are applicable to a manufacturing context in Southeast Asia or a healthcare setting in sub-Saharan Africa.

The transparency deficit as a systemic issue. The 82.9% non-reporting rate for country of data collection indicates that the transparency deficit is not marginal but systemic. The majority of primary studies in this corpus have been published in peer-reviewed venues without disclosing a piece of contextual information, namely the country in which data were collected, that is standard in other empirical social science fields and explicitly required by many journals and funding bodies. This is not a new problem: concerns about geographic reporting in behavioral and organizational research have been documented for decades. However, its manifestation in the human-AI collaboration literature, a field that prides itself on rigor and multidisciplinary breadth, is both surprising and consequential.

Several contributing factors may explain the non-reporting pattern. Authors publishing in information systems and management journals may assume that their institutional affiliation (typically visible in the article header) conveys sufficient geographic context, even when data collection took place in a different country. Authors conducting online survey research may consider the geographic context of respondents irrelevant when the study concerns attitudes or perceptions that may be universal. Journals and reviewers may not routinely require disclosure of data collection country as a condition of publication. Whatever the causes, the result is an evidence base whose geographic scope is largely opaque.

The sector non-reporting rate (30.6%), while substantially lower than the country rate, is also concerning. In some cases, sector non-reporting may reflect genuine difficulty in categorizing studies that cross sectoral boundaries or that use participant samples drawn from multiple industries. In other cases, it likely reflects the same reporting norm gap that produces country non-reporting: authors do not report what they are not required to report.

Structural causes of sectoral concentration. The extreme dominance of Business, Management and Finance (81.3% of usable sector observations) reflects structural features of the research community as much as features of the phenomenon itself. The major outlets for human-AI collaboration research, including journals in information systems, management science, and organizational behavior, are institutionally embedded in business schools and management faculties. The research community that produces this literature has strong

access to business and management settings and less systematic access to manufacturing shopfloors, clinical environments, military installations, or creative agencies. Funding patterns reinforce this concentration: grants from management research foundations and technology companies frequently target business applications of AI, while sector-specific AI research in healthcare, defense, and manufacturing tends to be conducted by separate research communities with different publishing traditions.

This structural explanation does not diminish the concern. It does, however, suggest that addressing sectoral concentration requires action at the structural level, including funding programs that explicitly incentivize sectoral diversity, interdisciplinary collaboration between management researchers and sector-specific experts, and journal editorial policies that encourage or require sectoral disclosure.

Analytical boundaries of this study. Several important caveats must be acknowledged. First, the concentration analysis is based only on the 17.1% of the corpus with usable country information and the 69.4% with usable sector information. If non-reporting is not random, the concentration estimates derived from the usable subset may not accurately represent the full corpus. For example, if Chinese and American researchers are more likely to report country of data collection than researchers from other countries, the observed concentration in China and the United States would overstate the true geographic concentration in the corpus. Conversely, if non-reporting is more common in studies from underrepresented countries, the true concentration may be even higher than observed. This uncertainty cannot be resolved without primary data from the study authors.

Second, the country and sector fields capture only two dimensions of the geographic and sectoral context of human-AI collaboration research. Other dimensions, including the size of organizations studied, the type of AI system deployed, the occupational characteristics of participants, and the regulatory environment, are not captured in this analysis. A comprehensive audit of contextual representativeness would require additional data extraction from primary sources.

Third, the sector categorization scheme, while developed systematically, inevitably involves judgment calls for ambiguous entries. The Business, Management and Finance category is intentionally broad and may absorb studies that could, with more detailed information, be assigned to narrower categories. Appendix B provides the full categorization mapping to allow readers to assess and, if desired, revise these assignments.

5 CONCLUSION

This paper has presented a systematic audit of geographic and sectoral reporting and distribution across 438 primary studies on human-AI team collaboration published between 2018 and 2026. The audit was motivated by two concerns: whether the transparency of the evidence base is sufficient to support geographic and sectoral inference, and, conditional on sufficient transparency, how concentrated the evidence base is across countries and sectors.

The results are unambiguous on the first question: the evidence base as a whole is not sufficiently transparent to support geographic inference for the majority of studies. Country of data collection is not usable in 82.9% of the corpus (363 of 438 studies). This is the primary empirical finding of the present analysis and should be treated as such. Any claim about the geographic distribution of human-AI collaboration research that does not acknowledge this non-reporting rate risks misrepresenting the scope of available evidence.

On the second question, the results among the studies that do disclose contextual information reveal pronounced concentration. Among the 75 studies with usable country information, China (38.7%) and the United States (17.3%) together account for 56.0% of usable observations. Sixteen distinct countries and one multi-country category are represented in the remaining 44.0%, but the majority of individual country representations consist of 1 or 2 studies, providing only sparse evidence about human-AI collaboration in those national contexts. Among the 304 studies with usable sector information, Business, Management and Finance accounts for 81.3%, with every other sector at 5.3% or below. Manufacturing and Industrial Production, Technology and Software Development, Creative Industries, and Military, Defense and Public Safety are each represented by 4 or fewer studies.

These findings carry three sets of implications.

For **researchers**, the results underscore the need to report the national context and organizational sector of data collection as standard practice, regardless of whether those details appear self-evident. Geographic and sectoral disclosure should be treated as a baseline transparency requirement, analogous to the reporting of sample size, study design, and analytic approach. Researchers working in underrepresented sectors and countries occupy a particularly valuable position for advancing the generalizability of the field.

For **journal editors and reviewers**, the results support the case for requiring disclosure of country of data collection and economic sector in the submission and review process. Given that 82.9% of existing studies did not disclose country information, editorial

requirements appear to be the most effective leverage point for improving reporting practice across the field.

For **funding bodies and grant agencies**, the results identify sectoral and geographic gaps that merit targeted investment. Research on human-AI collaboration in manufacturing settings in the Global South, in healthcare institutions in low-income countries, in creative and cultural industries across diverse national contexts, and in public-sector and military organizations is systematically underrepresented. Programs that specifically fund empirical work in these underrepresented contexts would contribute substantially to the cumulative validity of the evidence base.

This study has several limitations. First, because it operates entirely on metadata extracted from the source corpus, it cannot verify the accuracy of the original country and sector entries against the primary study texts. Normalization decisions are based on what was encoded in the database, and some entries may have been coded incorrectly in the original extraction. Second, the 82.9% country non-reporting rate means that concentration estimates based on the usable minority are uncertain in their representativeness of the full corpus. Third, the sector categorization scheme, while transparent and reproducible, involves judgment calls that other analysts might resolve differently, particularly for studies that span multiple sectors.

Future work should extend this audit in two directions. First, a targeted review of the non-reporting studies, drawing on full text access, could substantially reduce the pool of country-unknown records and provide a more accurate geographic picture. Second, longitudinal analysis of reporting rates over the 2018 to 2026 period would reveal whether transparency has improved over time and whether the field is moving toward the reporting norms that cumulative synthesis requires.

Human-AI collaboration is a field that aspires to produce practically applicable knowledge about how humans and artificial intelligence systems can work together effectively. Meeting that aspiration requires an evidence base that is not only large but geographically and sectorally diverse, and that is transparent enough to allow those dimensions to be assessed. The present study documents how far the field currently stands from that aspiration and establishes the descriptive baseline needed to track progress toward it.

ACKNOWLEDGMENTS

The author acknowledges the assistance of computational resources used in the normalization and analysis procedures described in Section 3. Statements regarding funding sources, competing interests, data and code availability, and pre-registration reflect the author's own determination and require confirmation by all contributing authors prior to submission.

REFERENCES

BOCKLISCH, F. *et al.* Hybrid decision-making in atmospheric plasma spraying enables human-machine teaming. The International Journal of Advanced Manufacturing Technology, v. 132, n. 9-10, p. 4941-4963, 2024. DOI: 10.1007/s00170-024-13595-8

BOSSEN, C.; PINE, K.H. Batman and Robin in healthcare knowledge work: human-AI collaboration by clinical documentation integrity specialists. ACM Transactions on Computer-Human Interaction, v. 30, n. 2, p. 1-29, 2023. DOI: 10.1145/3569892

DELL'ANNA, D. *et al.* From human teams to hybrid intelligence teams: identifying, characterizing, and evaluating foundational quality attributes. Autonomous Agents and Multi-Agent Systems, v. 40, n. 1, art. 10, 2026. DOI: 10.1007/s10458-025-09730-8

DELLERMANN, D. *et al.* Design principles for a hybrid intelligence decision support system for business model validation. Electronic Markets, v. 29, n. 3, p. 423-441, 2018. DOI: 10.1007/s12525-018-0309-2

ELLWART, T. *et al.* I vs. robot: sociodigital self-comparisons in hybrid teams from a theoretical, empirical, and practical perspective. Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO), v. 53, n. 3, p. 273-284, 2022. DOI: 10.1007/s11612-022-00638-5

EVERETT, S.S. *et al.* From tool to teammate in a randomized controlled trial of clinician-AI collaborative workflows for diagnosis. npj Digital Medicine, v. 9, n. 1, art. 409, 2026. DOI: 10.1038/s41746-026-02545-1

FABRI, L. *et al.* Disentangling human-AI hybrids. Business & Information Systems Engineering, v. 65, n. 6, p. 623-641, 2023. DOI: 10.1007/s12599-023-00810-1

FÜGENER, A. *et al.* Will humans-in-the-loop become Borgs? Merits and pitfalls of working with AI. MIS Quarterly, v. 45, n. 3, p. 1527-1556, 2021. DOI: 10.25300/MISQ/2021/16553

GEORGANTA, E.; PEUS, C.; NIESS, J. Interactive technologies through the lens of team effectiveness: an interdisciplinary systematic literature review. European Journal of Work and Organizational Psychology, v. 33, n. 2, p. 172-187, 2023. DOI: 10.1080/1359432X.2023.2178904

GEORGANTA, E.; ULFERT, A. Would you trust an AI team member? Team trust in human-AI teams. Journal of Occupational and Organizational Psychology, v. 97, n. 3, p. 1212-1241, 2024. DOI: 10.1111/joop.12504

HAGEMANN, V. *et al.* Human-AI teams: challenges for a team-centered AI at work. Frontiers in Artificial Intelligence, v. 6, art. 1252897, 2023. DOI: 10.3389/frai.2023.1252897

HENRICH, J.; HEINE, S.J.; NORENZAYAN, A. The weirdest people in the world. Behavioral and Brain Sciences, v. 33, n. 2-3, p. 61-83, 2010. DOI: 10.1017/S0140525X0999152X

HENRY, K.E. *et al.* Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. npj Digital Medicine, v. 5, n. 1, art. 97, 2022. DOI: 10.1038/s41746-022-00597-7

HU, W.; CHAN, C.K.Y. From user needs to AI solutions: a human-centered design approach for AI-powered virtual teamwork competency training. International Journal of Educational Technology in Higher Education, v. 22, n. 1, art. 52, 2025. DOI: 10.1186/s41239-025-00551-z

JIANG, J. *et al.* HAIGEN: towards human-AI collaboration for facilitating creativity and style generation in fashion design. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, v. 8, n. 3, p. 1-27, 2024. DOI: 10.1145/3678518

KRAUSE, F. *et al.* Managing human-AI collaborations within Industry 5.0 scenarios via knowledge graphs: key challenges and lessons learned. Frontiers in Artificial Intelligence, v. 7, art. 1247712, 2024. DOI: 10.3389/frai.2024.1247712

KRZYWDZINSKI, M. *et al.* How team organization influences the ability to solve automation failures: an experimental study on human-AI decision-making in teams. AI & SOCIETY, v. 41, n. 4, p. 3605-3620, 2025. DOI: 10.1007/s00146-025-02761-5

LEVINE, S.J.; WILLIAMS, B.C. Watching and acting together: concurrent plan recognition and adaptation for human-robot teams. Journal of Artificial Intelligence Research, v. 63, p. 281-359, 2018. DOI: 10.1613/jair.1.11243

LI, H.; TIAN, F. Advancing decision-making through AI-human collaboration: a systematic review and conceptual framework. Group Decision and Negotiation, v. 35, n. 2, art. 26, 2026. DOI: 10.1007/s10726-026-09980-1

MARRONE, R. *et al.* Understanding student perceptions of artificial intelligence as a teammate. Technology, Knowledge and Learning, v. 30, n. 3, p. 1847-1869, 2024. DOI: 10.1007/s10758-024-09780-z

NASER, M.Y.M.; BHATTACHARYA, S. Empowering human-AI teams via intentional behavioral synchrony. Frontiers in Neuroergonomics, v. 4, art. 1181827, 2023. DOI: 10.3389/fnrgo.2023.1181827

OFOSU-AMPONG, K.; RENNER-MICAH, A. From tool to teammate: a conjoined agency perspective on cross-cultural acceptability of AI as a co-worker. Discover Artificial Intelligence, v. 6, n. 1, art. 244, 2026. DOI: 10.1007/s44163-026-00998-7

ONTRUP, G. *et al.* Intelligent automated systems to support human teamwork: perceived invasiveness impacts team members' psychological reactions negatively. Cognition, Technology & Work, v. 27, n. 3, p. 591-609, 2025. DOI: 10.1007/s10111-025-00813-w

PAGE, M.J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, v. 372, art. n71, 2021. DOI: 10.1136/bmj.n71

PUERTA-BELDARRAIN, M. *et al.* A multifaceted vision of the human-AI collaboration: a comprehensive review. IEEE Access, v. 13, p. 29375-29405, 2025. DOI: 10.1109/ACCESS.2025.3536095

REZWANA, J.; MAHER, M.L. Designing creative AI partners with COFI: a framework for modeling interaction in human-AI co-creative systems. ACM Transactions on Computer-Human Interaction, v. 30, n. 5, p. 1-28, 2023. DOI: 10.1145/3519026

ROSENFELD, A.; RICHARDSON, A. Explainability in human-agent systems. Autonomous Agents and Multi-Agent Systems, v. 33, n. 6, p. 673-705, 2019. DOI: 10.1007/s10458-019-09408-y

SIEMON, D. Elaborating team roles for artificial intelligence-based teammates in human-AI collaboration. Group Decision and Negotiation, v. 31, n. 5, p. 871-912, 2022. DOI: 10.1007/s10726-022-09792-z

STIEGLITZ, S. *et al.* Collaborating with virtual assistants in organizations: analyzing social loafing tendencies and responsibility attribution. Information Systems Frontiers, v. 24, n. 3, p. 745-770, 2021. DOI: 10.1007/s10796-021-10201-0

VAN DER WAA, J. *et al.* Moral decision making in human-agent teams: human control and the role of explanations. Frontiers in Robotics and AI, v. 8, art. 640647, 2021. DOI: 10.3389/frobt.2021.640647

WANG, D. *et al.* Human-AI collaboration in data science. Proceedings of the ACM on Human-Computer Interaction, v. 3, n. CSCW, p. 1-24, 2019. DOI: 10.1145/3359313

WOOLLEY, A.W. Generative AI and collaboration: opportunities for cultivating collective intelligence. Journal of Organization Design, v. 15, n. 1, p. 45-50, 2025. DOI: 10.1007/s41469-025-00199-z

APPENDICES

APPENDIX A — COUNTRY NORMALIZATION MAPPING RULES

The following table documents the primary decision rules applied during country normalization. Raw field entries are grouped by the resolution they received. This documentation is provided to ensure full reproducibility of the normalization procedure and to allow other researchers to assess or revise individual decisions.

Table A1

Representative raw entries by normalization category

Normalized Category	Representative Raw Entry Patterns
Not reported	Blank field; "não informado"; "não especificado"; "não reportado"; unresolvable descriptive text; entries providing contextual clues but not confirming data collection location (e.g., HIPAA reference suggesting US context)
Not applicable	"não aplicável; corpus internacional indexado na Web of Science"; entries referring to multi-country index corpora or theoretical studies with no empirical data collection country
Multiple countries	Comma-separated lists of 2 or more country names; "80 países (incluindo EUA, Brasil, Islândia...)"; "múltiplos países, não especificado individualmente"; "não aplicável/múltiplos"; country lists such as "Alemanha, Suíça e Reino Unido"; "Estados Unidos e Coreia do Sul"; "Malásia, Taiwan e Estados Unidos"
China	"China"; "chineses" (as participant attribute, e.g., "profissionais chineses")
United States	"Estados Unidos"; "EUA"; "United States"; "US"
Germany	"Alemanha"; "Germany"
Japan	"Japão"; "Japan"
Australia	"Austrália"; "Australia"
Netherlands	"Países Baixos"; "Netherlands"; "Holanda"
United Kingdom	"Reino Unido"; "UK"; "United Kingdom"
India	"Índia"; "India"
Jordan	"Jordânia"; "Jordan"
Ecuador	"Equador"; "Ecuador"

Normalized Category	Representative Raw Entry Patterns
Ireland	"Irlanda"; "Ireland"
Denmark	"Dinamarca"; "Denmark"
Finland	"Finlândia"; "Finland"
Kosovo	"Kosovo"
Colombia	"Colômbia"; "Colombia"
Kuwait	"Kuwait"

Source: Prepared by the authors.

Resolution notes: CSIRO (Australian national research agency) appeared in a field otherwise flagged as unspecified; the entry was retained as *Not reported* because the institution's location does not confirm the country of data collection. All entries containing “China” as part of compound phrases (e.g., “Chinese professionals”) were resolved to China where the phrase unambiguously identified participant nationality.

APPENDIX B — SECTOR CATEGORIZATION SCHEME

The eight-category scheme applied in this study is described in detail in Section 3.3 of the main text. This appendix provides illustrative mapping examples and clarifies priority rules for ambiguous cases.

Table B1

Sector categories with illustrative descriptors and priority order

Priority	Category	Illustrative Descriptors
1	Healthcare and Medicine	Clinical decision support; diagnostic AI; electronic health records; nursing; medical imaging; hospital operations; patient monitoring
2	Military, Defense and Public Safety	Unmanned vehicle operations; combat simulation; law enforcement; emergency response; cybersecurity operations
3	Manufacturing and Industrial Production	Assembly lines; quality control; robotic welding; logistics; industrial automation; supply chain operations
4	Technology and Software Development	AI-assisted programming; code review; software design; system architecture; LLM-assisted development
5	Education and Academia	AI tutoring; student-AI collaboration; academic group work; intelligent learning systems; higher education
6	Creative Industries	Fashion design; music composition; visual art; journalism; brand management; game design
7	Experimental or Generic Task	Laboratory decision tasks; generic cognitive tasks; abstract group problem-solving; behavioral experiments with no sector designation
8	Business, Management and Finance	Corporate strategy; financial forecasting; HR management; marketing; consulting; organizational decision-making; generic managerial contexts
—	Not reported	No sector information provided; blank; "not specified"; "not applicable" without further context

Source: Prepared by the authors.

Priority application: When a study's description contained elements matching multiple categories, the higher-priority category was assigned. For example, a study on AI in hospital supply chain management was assigned to Healthcare and Medicine (priority 1) rather than Manufacturing (priority 3) or Business/Management (priority 8). A study on AI in software security in a defense context was assigned to Military, Defense and Public Safety (priority 2) rather than Technology and Software Development (priority 4).

The Business, Management and Finance category serves as the broadest residual category for organizational research. Studies describing their context in generic organizational or managerial terms, without a more specific sectoral anchor, default to this

category. Its high frequency in the corpus (247 of 304 usable observations, 81.3%) reflects both genuine disciplinary concentration in business research and the breadth of the category definition.

Manuscript prepared in accordance with the editorial guidelines of Revista ReGeo (ISSN 2177-3246). Times New Roman 12pt, 1.5 line spacing. ABNT NBR 14724, NBR 10520, NBR 6023.