CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Unpacking human–AI collaboration: conceptualisations and emerging research streams
Authors	Wang, Yu;Heavin, Ciara;de Paula, Danielly
Publication date	2026-04-16
Original Citation	Wang, Y, Heavin, C & de Paula, D 2026, 'Unpacking human–AI collaboration: conceptualisations and emerging research streams', Journal of Decision Systems, vol. 35, no. 1, 2653697. https://doi.org/10.1080/12460125.2026.2653697
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://www.scopus.com/pages/publications/105036028797 - 10.1080/12460125.2026.2653697https://www.tandfonline.com/doi/pdf/10.1080/12460125.2026.2653697
Rights	© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is notaltered, transformed, or built upon in any way. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-30 03:45:37
Item downloaded from	https://hdl.handle.net/10468/18842

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh



Journal of Decision Systems

ISSN: 1246-0125 (Print) 2116-7052 (Online) Journal homepage: www.tandfonline.com/journals/tjds20

Unpacking human–AI collaboration: conceptualisations and emerging research streams

Yu Wang, Danielly de Paula & Ciara Heavin

To cite this article: Yu Wang, Danielly de Paula & Ciara Heavin (2026) Unpacking human–AI collaboration: conceptualisations and emerging research streams, Journal of Decision Systems, 35:1, 2653697, DOI: 10.1080/12460125.2026.2653697

To link to this article: https://doi.org/10.1080/12460125.2026.2653697

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 16 Apr 2026.

Submit your article to this journal

Article views: 180

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tjds20

JOURNAL OF DECISION SYSTEMS
2026, VOL. 35, NO. 1, 2653697
https://doi.org/10.1080/12460125.2026.2653697

Taylor & Francis
by informa

OPEN ACCESS Check for updates

Unpacking human–AI collaboration: conceptualisations and emerging research streams

Yu Wang, Danielly de Paula and Ciara Heavin

Department of Business Information Systems, Cork University Business School, University College Cork, Cork, Ireland

ABSTRACT

Despite the growing adoption of AI across organisations, fragmented knowledge of key concepts and measurement approaches constrains systematic evaluation and cumulative progress. This review aims to harmonise the conceptual vocabulary of HAIC, map how conceptual framings align with evaluation emphases, and surface an stream structure from co-occurrence patterns, thereby supporting more comparable HAIC assessment designs. We reviewed 149 peer-reviewed studies (2018–2025) from AISeL, ACM, and IEEE using a concept-centric approach. We identified 16 recurring themes for coding and mapping, 10 of which we further defined by synthesising consistent definitional passages to specify boundary conditions that distinguish collaboration from tool use and automation. To identify concept–evaluation link patterns, we computed a co-occurrence matrix and a Jaccard-normalised overlap indicator. Our findings reveal four preliminary streams that organise these linkages. Overall, the study proposes a harmonised cross-disciplinary vocabulary and an empirically grounded mapping that strengthens interpretability and comparability in HAIC assessment.

ARTICLE HISTORY
Received 17 February 2026
Accepted 26 March 2026

KEYWORDS
Human–AI collaboration; human–AI teaming; hybrid intelligence; intelligence augmentation; human-AI decision making

1. Introduction

With the AI Index Report 2025 highlighting that 78% of surveyed organisations used AI in 2024, up from 55% in 2023 (AI Index Steering Committee [AI], 2025), organisational uptake of AI has accelerated. Human – AI collaboration (HAIC) is becoming a mainstream way of organising analytical, decision-making, and creative work (Berretta et al., 2023), leaving many organisations with a practical question of how to combine human judgement with model capability in everyday work (Wang et al., 2023). As AI becomes embedded in organisational work, collaboration quality often becomes a design and governance concern rather than only a usability issue (Chen et al., 2024), especially when AI adds value by complementing human judgement under uncertainty and equivocality rather than replacing it outright (Vaccaro et al., 2024). Organisations therefore need to align model capability with human responsibility, particularly when errors carry material or reputational consequences (Shneiderman, 2020), and interaction research reinforces this point by showing that reliance on AI unfolds through feedback loops shaped by transparency, controllability, and expectation management (Amershi et al., 2019). Building on this reliance-focused view, similar concerns are prominent in AI-based decision support systems (DSS) that increasingly mediate organisational decision-making: trust, explainability, and transparency shape how users accept, contest, and evaluate AI-supported recommendations. Recent work on generative AI decision-making systems further suggests that transparency and ethical alignment are becoming more salient design and evaluation concerns (Goktas, 2024), closely echoing HAIC debates on calibrated reliance, explanatory support, and governance mechanisms (Saup et al., 2026).

Despite rapid growth, the HAIC literature has not fully converged on a shared conceptual vocabulary (Berretta et al., 2023). Researchers describe closely related phenomena using different labels and attach different scope conditions (Fabri et al., 2023), which complicates cumulative knowledge building because studies may appear comparable while referring to different collaboration arrangements (Gomez et al., 2025). When concepts vary, evaluation choices become harder to interpret because researchers lack a stable

CONTACT Ciara Heavin c.heavin@ucc.ie Cork University Business School, University College Cork, Room 3.85, O'Rahilly Building, College Road, Cork T12 CY82, Ireland

Supplemental data for this article can be accessed online at https://doi.org/10.1080/12460125.2026.2653697

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

reference point for what counts as collaboration success (Puerta-Beldarrain et al., 2025). Evaluation practices also appear fragmented (Lai et al., 2023): studies propose diverse metrics but often do not explain how measurement choices follow from the specific conceptual framing of collaboration (Fragiadakis et al., 2024; Vereschak et al., 2021). A cross-disciplinary synthesis is therefore needed to clarify recurring concepts and boundary conditions, and to map how concept evaluation linkages co-occur across the literature.

HAIC is inherently interdisciplinary and has developed in parallel across IS and CS. IS research emphasises socio-technical conceptualisations and organisational implications (Sawyer & Jarrahi, 2014), whereas CS research contributes operationalisations, task paradigms, and evaluation approaches (Wegner, 1976). We therefore integrate IS and CS outlets to reduce discipline-specific blind spots and to enable an interpretable mapping of how conceptual and evaluation themes co-occur across the corpus. While our review integrates both IS and CS outlets, the retrieved corpus is predominantly composed of CS studies; in pre-tests, the IS – CS gap was much smaller when searching HAIC terms alone, but widened once the conceptualisation/evaluation block was added. This highlights a gap in the IS literature, where evaluation-focused research on human – AI collaboration remains underdeveloped, particularly when 'evaluation' and 'conceptualisation' are used as keywords. We therefore interpret cross-disciplinary patterns as indicative rather than as balanced population estimates. To address this, we propose the following research objective (RO) and questions (RQ):

RO: To synthesise and organise IS and CS research on HAIC conceptualisation by consolidating recurring concepts and definition patterns, and by mapping theme co-occurrence to derive a preliminary set of research streams.

Main RQ: *How do IS and CS studies conceptualise HAIC, and what theme relatedness patterns structure this literature?*

Sub-RQ1: *What core HAIC concepts and definition patterns recur across IS and CS studies, and what boundary conditions distinguish them?*

Sub-RQ2: *What theme co-occurrence patterns emerge across the corpus, and how do these patterns cluster into preliminary research streams?*

To address these research questions, we reviewed 149 HAIC studies published between 2018 and 2025 across IS and CS outlets. We developed a concept table to consolidate definitional evidence and boundary conditions, and we mapped cross-theme relatedness using a theme co-occurrence matrix with a normalised overlap indicator. These results provide the basis for an indicative stream structure reported in the Results.

2. Conceptual background

Human – AI collaboration is better treated as a family of related concepts than as a single settled construct (Berretta et al., 2023). Across studies, authors variously foreground AI as a teammate with autonomy (Berretta et al., 2023), purposeful integration of complementary capabilities (Dellermann et al., 2019), or augmentation (Zhou et al., 2021) that enhances human capability rather than replaces it. Autonomy focused work similarly defines human autonomy teams as configurations where humans and autonomous systems pursue a shared goal (Krausman et al., 2022). We treat these conceptual choices as boundary conditions when synthesising definitions and comparing findings.

A second line of work shifts from defining collaboration to specifying collaboration modes and governance mechanisms. Collaboration is often framed as interdependence and joint contribution in problem-solving, which separates collaboration from one way tool use (Metcalfe et al., 2021). Mechanisms such as agency and control allocation (Dubey et al., 2020), delegation and task allocation (Spitzer et al., 2025), and mixed initiative interaction (Li et al., 2025) then become central because they shape who initiates, decides, and executes during a task. Evaluation becomes fragile when studies report outcomes without stating the assumed mode, or when metrics travel across settings with different authority structures and interaction dynamics (Fragiadakis et al., 2024). This challenge is reinforced by HAIC's cross disciplinary roots, where IS foregrounds socio technical embedding (Sawyer & Jarrahi, 2014) and CS and HCI foreground task paradigms

and interaction mechanisms (Amershi et al., 2019). Accordingly, we treat conceptual consolidation and concept to evaluation mapping as linked steps in the synthesis.

3. Methodology

We conducted a structured literature review across authoritative IS and CS outlets: AIS eLibrary (AISeL), ACM Digital Library, and IEEE Xplore (Webster & Watson, 2002). We used two keyword blocks, combining HAIC-related terms (e.g. human – AI collaboration/teaming, hybrid intelligence, AI augmentation) with conceptualisation and evaluation terms (e.g. definition, framework, taxonomy, evaluation, measurement, metrics) using AND. We applied the same query logic across all three databases, and the corpus was restricted to peer-reviewed journal and conference publications from 2018 to 2025. We set 2018 as the lower bound because work from this period consistently frames human and AI as interdependent collaborators and foregrounds interaction and evaluation considerations aligned with our review questions (Jarrahi, 2018), while 2025 was used as the upper bound to capture the most recent publication cycle available at screening.

3.1. Screening procedure

To ensure transparency and reproducibility, we summarise the search and screening workflow and record counts in Figure 1.

The search returned 184 records from AISeL, 1,081 from ACM, and 936 from IEEe. Because we combined HAIC terms with a conceptualisation and evaluation keyword block, the retrieved records were skewed towards CS outlets. In pre-tests using HAIC terms only, the IS – CS gap was much smaller, but the IS yield dropped sharply once conceptualisation and evaluation terms were added, suggesting that up to recently fewer IS papers foreground these aspects in searchable fields. After deduplication, 2,193 unique records remained. We applied a two-stage screening process (title/abstract, then full text) using predefined inclusion and exclusion criteriA. We retained studies that explicitly conceptualise HAIC (e.g. provide definitional passages or describe collaboration forms) and/or operationalise and evaluate HAIC (e.g. measures of collaboration processes, quality, or outcomes). We excluded papers that treated 'collaboration' only as background context or focused primarily on

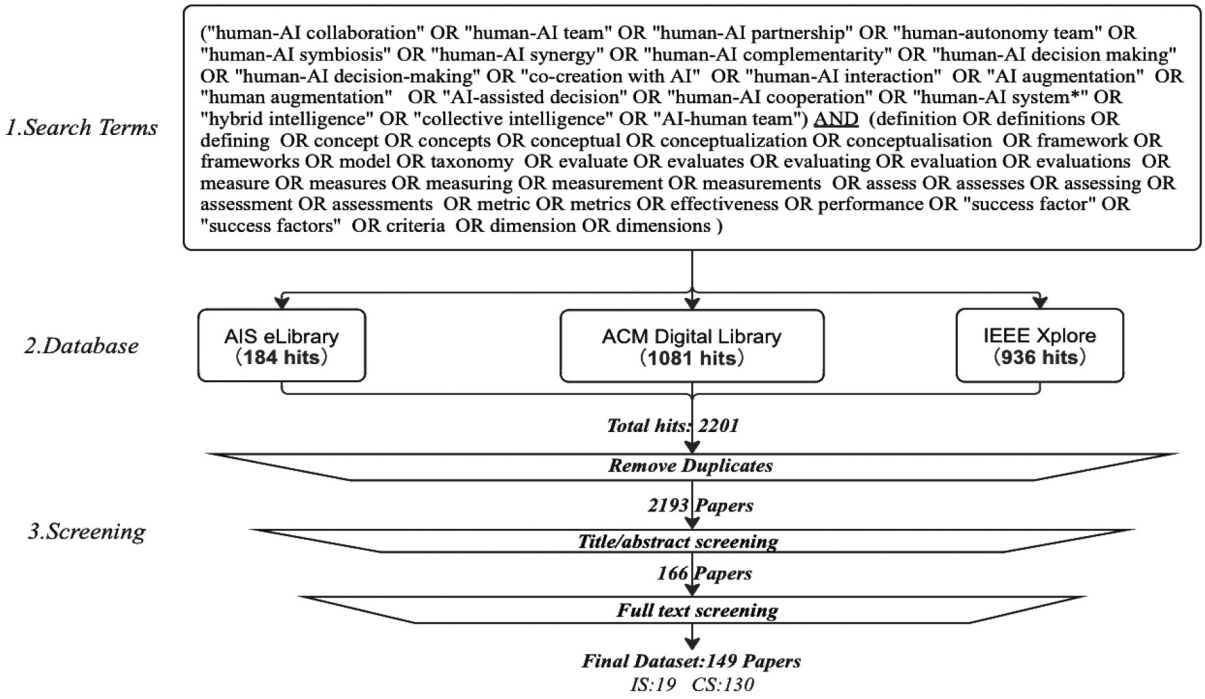


Figure 1. Procedure for extracting and screening literature.

technical development without evaluating the human – AI collaboration phenomenon (e.g. Zabaleta et al., 2019). Following full-text screening, 149 studies were retained; Figure 1 summarises the search and screening process. All analyses in this paper are based on this fixed corpus of 149 studies published up to 2025. We may extend the review in future work through backward and forward citation searching, any additional studies published in 2026 fall outside the current search window.

3.2. Data extraction & preliminary coding scheme

We extracted (1) descriptive metadata, including publication year, venue or discipline, and task context, and (2) analytic data aligned with our two main synthesis outputs, namely a concept table (see Table 1) and a theme co-occurrence matrix (see Table 2). For concept synthesis, we extracted definition related passages from each study and treated these excerpts as evidence. We then triangulated two to three excerpts per concept to formulate an integrative definition and specify boundary conditions that distinguish the concept from adjacent terms. Table 1 reports ten concepts because these were the themes for which the included studies most consistently provided explicit definitional passages and distinguishable boundary conditions; for instance, definitional passages typically use explicit 'refers to … ' formulations that specify the collaboration relationship, whereas boundary conditions explicitly exclude adjacent arrangements such as one-way tool use or automation-first substitution logics. We retained a broader theme dictionary of 16 themes for corpus level coding and co-occurrence mapping. For matrix construction, we applied the 16 theme dictionary and coded each paper for the presence (1) or absence (0) of each theme. Coding was led by the first author and refined through pilot coding and iterative calibration discussions with the co-authors, during which overlapping labels were merged and ambiguous cases were resolved to support consistent corpus-level coding. Theme co-occurrence was computed as the number of studies in which two themes were both coded as present and summarised in a co-occurrence matrix. To interpret overlap relative to theme prevalence, we also computed the Jaccard index for each theme pair (Jaccard, 1901), which helps reduce the risk of treating highly prevalent themes as central solely because they appear often (Fletcher & Islam, 2018).

3.3. Synthesis strategy

Our synthesis proceeds in three steps. First, we develop a concept table by synthesising definitional evidence and recurring boundary conditions across the included studies (see Table 1). Second, we use the theme co-occurrence structure to summarise how themes relate across studies (see Table 2). Third, we translate salient co-occurrence patterns into a set of preliminary research streams by starting from anchor theme pairs, expanding each stream with closely connected neighbouring themes, and corroborating the interpretation through targeted close reading. To reduce the risk that highly prevalent themes dominate this grouping, we interpret patterns in light of theme prevalence and treat hub-like themes as background rather than as stream drivers. The resulting streams serve as an organising lens for reporting and discussing the results, not as exhaustive categories or causal claims, and exemplar studies are selected to ground each stream's narrative.

4. Results

This section reports the results of our concept-centric synthesis of the 149 included studies, focusing on how HAIC is conceptualised across IS and CS and how these conceptualisations connect to evaluation and governance related themes. Table 1 presents the 16 concepts identified in the review and reports synthesised definitions only where sufficiently consistent definitional passages were available, resulting in 10 defined concepts (A – J). Table 2 reports the co-occurrence structure across the coded themes in the corpus, which we use to quantitatively identify robust link patterns and to derive an indicative set of preliminary research streams.

Table 1. HAIC key concepts, definitions and distinguishing characteristic.

	Theme name	Concept type	Representative definition (synthesised)	Key characteristics/boundaries
A	Human-AI Collaboration/ Interworking	Collaboration definition	Human–AI collaboration/interworking refers to a collaborative relationship in which humans and AI (or AI-enabled systems) engage in complementary division of labour and mutual interdependence to solve problems or accomplish tasks Fabri et al. (2023). Its core lies in a collaborative process of joint output and/or joint decision-making, rather than a one-way "human uses AI as a tool" mode of use Metcalfe et al. (2021).	•Joint participation in problem solving and task accomplishment Fabri et al. (2023). •Close interworking characterised by tight human–AI coupling and mutual dependence Fabri et al. (2023). Boundary: Excludes "Human or AI" framings where AI is merely a tool or alternative rather than a collaborator Metcalfe et al. (2021).
B	Human-AI Teaming	Collaboration definition	Human–AI teaming refers to a team composed of at least one human member and at least one AI system/agent with some degree of autonomy, who collaborate toward a shared task or goal Krausman et al. (2022).	•A team consists of one or more humans and one or more autonomous systems or intelligent agents collaborating toward a task or goal Krausman et al. (2022). •AI is treated as a teammate, rather than merely a tool or a back-end automation module Jha et al. (2024). Boundary: Excludes tool-like or back-end automation roles that do not qualify as team membership Schelble et al. (2022).
C	Hybrid Intelligence	Collaboration definition	Hybrid Intelligence (HI) refers to the intentional integration of human intelligence and AI/ machine intelligence Schlup and Corradini (2024), whereby interaction and communication between human and machine actors enable them to jointly address cognitively intensive or complex tasks Rockbach et al. (2022) and achieve outcomes that are difficult to attain with humans or AI alone Krinkin et al. (2021).	•Combines human and machine intelligence to address cognition-intensive tasks Schlup & Corradini (2024). •Integration is realised through interaction and communication between human and artificial agents Rockbach et al. (2022). Boundary: Requires complementarity such that goals are unattainable by either humans or machines alone Krinkin et al. (2021).
D	Intelligence Augmentation	Collaboration definition	Intelligence Augmentation (IA) refers to the use of information technologies and AI to enhance human capabilities, intelligence, and performance. Its central aim is to improve what humans can do and how well they can do it Zhou et al. (2021), rather than to replace humans through automation Paul et al. (2022).	•Explicitly oriented toward human capability uplift rather than machine-led autonomy Paul et al. (2022). Boundary: Excludes automation-first substitution logics, instead framing IA as improving human capacity and quality of life Zhou et al. (2023).
E	Agency & Control Allocation	Collaboration mode definition	Agency & Control Allocation refers to the distribution and dynamic adjustment of task initiation, decision authority, action execution, and ultimate accountability between human actors and AI or autonomous systems in human–AI collaboration or teaming contexts Van der Aalst (2021).	•Control and contribution to decision making can shift between user and system Dubey et al. (2020). Boundary: Focuses on how to keep humans in control when deciding what and how to automate Van der Aalst (2021).
F	Delegation/Task Allocation	Collaboration mode definition	Delegation or task allocation refers to the mechanism through which specific task instances, subtasks, or decision authority are assigned or transferred between humans and AI in human–AI collaboration Spitzer et al. (2025).	•Its purpose is to improve team-level task performance, such as accuracy and efficiency Spitzer et al. (2025) •Allocation decisions are typically informed by considerations of complementary capabilities, contextual information, uncertainty or confidence, and expected performance Fuchs et al. (2024). Boundary: Includes dynamic delegation of decisions to the actor deemed more suitable at a given time Fuchs et al. (2024).
G	Mixed-Initiative Interaction	Collaboration mode definition	Mixed-initiative interaction refers to a human–AI collaborative interaction paradigm in which both humans and AI can initiate, advance, or take over actions, decisions, or creative contributions during task execution Li et al. (2025).	•Humans and AI jointly contribute to decisions, with AI treated as a peer collaborator Li et al. (2025). •Initiative dynamically shifts and is exchanged during complex work processes Muller and Weisz (2022). Boundary: Requires bilateral initiative, including contexts where both parties take initiative in co-creation Rezwana and Maher (2023).
H	Trust Calibration/ Reliance Dynamics	Collaboration mode definition	Trust calibration and reliance dynamics refer to the degree to which a person's trust in and reliance on AI, in collaboration or teaming contexts, align with the AI's actual capabilities and performance, and how this alignment evolves over time Duan et al. (2024).	•Appropriate calibration requires matching trust and reliance to true system capability (Pareek et al. (2024). Boundary: Emphasises case-specific decisions about when to trust or distrust in use Zhang et al. (2020).

(*Continued*)

Table 1. (Continued).

	Theme name	Concept type	Representative definition (synthesised)	Key characteristics/boundaries
I	Explainability/ Transparency (mechanism-linked)	Collaboration mode definition	Explainability and transparency refer to providing interpretable and actionable explanatory information that makes an AI system's reasoning basis, problem identification logic, prediction rationale, and relevant uncertainty or behavioural grounds transparent to human collaborators Kim et al. (2023), thereby directly supporting collaborative judgment and action in specific task contexts Zhang et al. (2020).	•Makes AI more understandable to people through interpretable explanations Kim et al. (2023). •Prioritises practically useful information that improves collaboration rather than technical detail Kim et al. (2023).
J	Accountability/ Responsibility (decision-chain-linked)	Collaboration mode definition	Accountability and responsibility refer to the explicit and traceable assignment of responsibility for decision processes and outcome consequences in contexts where humans and AI jointly contribute to decisions or task execution Zhang et al. (2020).	•Addresses who is responsible when humans and AI jointly shape decisions and outcomes Muller and Weisz (2022). •Requires explicit responsibility assignment across recommendation, acceptance or override, and outcomes Zhang et al. (2020).
K	Evaluation, Measurement		Not synthesised	Used as coding theme, not defined
L	Performance, Effectiveness		Not synthesised	Used as coding theme, not defined
M	Productivity/Work Outcomes		Not synthesised	Used as coding theme, not defined
N	Communication/Coordination		Not synthesised	Used as coding theme, not defined
O	Autonomy Level		Not synthesised	Used as coding theme, not defined
P	Framework, Conceptual Model		Not synthesised	Used as coding theme, not defined

Table 2. Theme co-occurrence counts & Jaccard matrix.

	K	E	P	H	I	A	F	G	N	M	L	J	C	D	O
K		**30**	**32**	**27**	**28**	**25**	**21**	16	**17**	12	**38**	12	11	10	4
E	0.27		16	17	16	**21**	13	13	12	11	3	10	13	9	7
P	0.28	0.21		13	12	16	15	16	13	12	10	11	11	9	3
H	0.25	0.26	0.18		**24**	16	12	10	12	13	8	11	10	8	3
I	0.27	0.26	0.17	**0.49**		15	13	11	12	11	9	11	6	10	2
A	0.23	0.38	0.24	0.28	0.28		12	11	12	10	1	10	10	7	0
F	0.21	0.24	0.26	0.23	0.28	0.26		14	12	12	9	11	9	5	2
G	0.16	0.26	0.31	0.20	0.26	0.26	**0.45**		13	10	7	10	7	7	3
N	0.17	0.24	0.24	0.26	0.29	0.29	0.38	0.50		10	2	11	5	5	2
M	0.12	0.22	0.23	0.30	0.28	0.24	0.40	0.37	0.38		2	10	8	6	1
L	0.37	0.04	0.12	0.11	0.13	0.01	0.15	0.13	0.03	0.03		1	2	1	2
J	0.12	0.22	0.22	0.27	0.31	0.27	0.41	0.43	0.52	0.50	0.02		5	6	0
C	0.10	0.23	0.17	0.18	0.11	0.20	0.21	0.18	0.12	0.22	0.03	0.14		0	4
D	0.09	0.17	0.16	0.16	0.24	0.16	0.13	0.23	0.16	0.21	0.02	0.24	0.00		3
O	0.04	0.15	0.05	0.07	0.05	0.00	0.06	0.11	0.07	0.04	0.04	0.00	0.12	0.12	

Note: **Upper triangle** (white background) = co-occurrence counts; **Lower triangle** (light-grey background) = Jaccard. **Bold (upper)** = Top10 counts among pairs; **Bold (lower)** = pairs with Jaccard ≥ 0.45 AND n ≥ 14. The union of these bolded pairs constitutes the core link set used to derive the preliminary stream structure. All statistics are computed after excluding Theme B (Human–AI Teaming).

4.1. Most prominent conceptualisations of human – AI collaboration in IS and CS

Across the cross-disciplinary corpus, we identified 16 recurring concepts (Table 1). These concepts span both what collaboration is and how collaboration is enacted and governed. Four concepts (A – D) primarily define collaboration itself by specifying the relationship between human and AI actors (e.g. interdependence, joint contribution, and the intended purpose of combining capabilities). A further six concepts (E – J) define collaboration in terms of modes and mechanisms that structure interaction and responsibility, such as who holds agency, how control is allocated, how tasks are delegated, how initiative is shared, and how reliance is managed. The remaining six themes (K – P) function mainly as coding themes that capture evaluation dimensions, outcomes, or analytic lenses; these were retained to support cross-study mapping but were not treated as definitional concepts because explicit definitional passages were not consistently available.

In this section, the 'most prominent conceptualisations' refer to the conceptual anchors identified by (i) recurring presence across IS and CS studies and (ii) sufficiently consistent definitional passages to support synthesis.On this basis, the ten synthesised concepts (A – J) represent the most stable conceptual anchors in the corpus. Within the core collaboration definitions (A – D), studies converge on the idea that HAIC involves more than one-way tool use by emphasising some form of joint contribution or interdependence, while differing in how strongly the AI is treated as an autonomous partner and what the human – AI arrangement

is meant to achieve. For example, teaming-oriented accounts foreground role expectations and coordination around a shared task, whereas hybrid-intelligence and augmentation framings emphasise complementary strengths but vary in whether the emphasis is on joint system-level capability versus enhancing human capability.

The mode/mechanism concepts (E – J) shift attention from 'what HAIC is' to 'how it operates'. Here, collaboration is conceptualised through arrangements, whether explicitly designed or emergent in practice, that shape how authority, initiative, and accountability are distributed across a task. This layer makes the evaluation problem more explicit: what counts as 'good collaboration' depends not only on outcomes but also on whether governance, reliance, and explanation mechanisms support appropriate human control and responsibility in the interaction. Taken together, Table 1 shows that conceptualisations in IS and CS are not competing alternatives but often complementary lenses: core definitions establish the collaboration relationship, while mode/mechanism concepts specify the interactional and governance structures through which that relationship is realised.

4.2. Theme co-occurrence and concept – evaluation linkages

To examine how conceptual and evaluation themes connect across studies, we constructed a theme co-occurrence matrix and a normalised association indicator (Table 2). Co-occurrence counts capture absolute overlap: for each theme pair, the count reports how many studies coded both themes as present. Because some themes are broadly used across the corpus, raw counts can overstate apparent 'centrality' for high-prevalence themes. We also report the Jaccard index, which expresses overlap relative to each theme's overall prevalence and helps distinguish selective coupling from popularity alone.

To identify robust linkages and reduce hub-driven effects, we re-ranked theme pairs after excluding Theme B (Human – AI Teaming), which is the most prevalent theme in the corpus and co-occurs with many other themes. We compute both the co-occurrence counts and the Jaccard similarity on the B-excluded matrix so that the strongest links are not dominated by a single high-prevalence theme. While other normalisation strategies are possible; we adopt this procedure because it is transparent, easy to reproduce, and helps surface more selective couplings among the remaining themes.We then selected candidate links using a two-part rule. First, we retained the Top 10 pairs by co-occurrence count to capture the most frequently co-discussed connections (the smallest count in this set was $n = 17$). Second, to avoid missing tightly coupled but less frequent mechanism pairs, we added the highest-frequency pair (s) with Jaccard $\geq$ 0.45 that still met a minimum co-occurrence threshold ($n \geq 14$). After removing any duplicates across the two steps, the resulting link set (bolded and shaded grey in Table 2) provides the basis for grouping recurring patterns into preliminary research streams. Importantly, co-occurrence indicates association rather than causality; accordingly, we treat these streams as organising lenses, and we corroborate their interpretation via targeted close reading of the main contributing studies.

To make the stream derivation transparent, we constructed streams from the selected link set in three steps. First, we treated each stream's anchor pair as a seed and assigned it to a provisional cluster (Stream A: Measurement & Performance; Stream B: Governance & Control; Stream C: Trust Calibration; Stream D: Delegation & Mixed-Initiative). Second, we expanded each cluster by adding any theme that (i) appears in the selected link set and (ii) connects directly to the cluster through at least one retained pair; where a theme connected to multiple clusters, we assigned it to the cluster with the stronger and more coherent set of retained links, and treated the remaining connections as cross-stream ties. Third, we corroborated the resulting cluster labels through targeted close reading of the main contributing papers for each anchor pair to ensure that the stream name reflects the shared conceptual and evaluative focus indicated by the link patterns.

First, the Measurement & Performance stream is signalled by links between evaluation/metrics themes and outcome-oriented themes (e.g. performance and effectiveness, and where relevant productivity/work outcomes). Studies in this cluster commonly treat collaboration quality as demonstrable gains in task success, accuracy, or efficiency under specific task paradigms, often making evaluation design and metric choice central to the contribution (He et al., 2023).

Second, the Governance & Control stream is reflected in repeated associations among themes that specify how authority and responsibility are distributed (e.g. agency and control allocation, accountability/

Table 3. Summary of preliminary HAIC research streams derived from theme Co-occurrence.

Research stream	IS/CS count	Key Ideas	Representative studies	Future research
Stream A: Measurement and Performance	6/42	• Metric-driven evaluation of HAIC; • Short-horizon, ground-truth task bias evaluation paradigms; • Uneven construct-to-metric alignment when metrics are weakly anchored to frameworks/definitions	He et al. (2023) Jha et al. (2024). Sadeghian et al. (2024)	• Integrate work-level consequences (e.g. experience) alongside performance when claiming organisational relevance •Test robustness of concept–metric mappings
Stream B: Governance & Control	6/33	• Governance treated as both a design object and an evaluative concern; • Allocation of authority and responsibility under uncertainty; • Governance salience in high-stakes contexts;	Dubey et al. (2020) Teodorescu et al. (2021) Caldwell et al. (2022)	• Test whether control-allocation strategies remain robust as tasks evolve, uncertainty rises, or incentives shift. • Assess governance effects on work-level outcomes (eg., meaningfulness)
Stream C: Trust Calibration & Explainability	2/39	• Frame calibration as aligning reliance with actual capability (not maximising trust). • Explanations as case-based reliability cues, not generic transparency. • Calibration is dynamic over time with feedback.	Zhang et al. (2020) Kim et al. (2023) Duan et al. (2024) Pareek et al. (2024)	• Specify and test mechanism pathways: explanation design → reliance behaviour → calibration (incl. post-error repair). • Measure calibration outcomes directly (not only downstream accuracy).
Stream D: Delegation & Mixed Initiative	3/13	• Task ownership and handoffs; • Collaboration quality as experienced in handoffs; • Especially salient in open-ended/co-creative tasks;	Hemmer et al. (2023) Pinski et al. (2023) Rezwana and Maher (2023) Muller and Weisz (2022)	• Boundary conditions for handoff intensity; • Directionality and mechanism tests;

Note: Counts are non-exclusive; one study may contribute to multiple streams. IS/CS is determined by database source (AISeL = IS; ACM/IEEE = CS).

responsibility, and autonomy-related considerations) and their connections to evaluation. Here, collaboration quality is often framed through how systems structure decision rights, oversight, and responsibility boundaries, especially where risk, uncertainty, or organisational consequences make governance salient (Dubey et al., 2020).

Third, the Trust Calibration stream is characterised by selective coupling between reliance dynamics and informational support themes such as explainability/transparency. In this cluster, collaboration quality is commonly treated as appropriate reliance over time and across cases, with explanatory support frequently co-occurring as part of how calibration is examined in evaluation designs (Kim et al., 2023). As with all co-occurrence findings, these patterns reflect association rather than evidence of causal effects.

Finally, the Delegation & Mixed-Initiative stream is best read as an interaction-mechanism cluster rather than a fully independent agenda: the core pair captures how collaboration is enacted through task routing and initiative sharing. Its connections to governance and trust-related themes suggest that delegation and initiative often function as the 'plumbing' that operationalises control allocation and shapes reliance dynamics in practice, instead of appearing only as implementation detail (Hemmer et al., 2023).

Table 3 summarises these four streams by consolidating their key ideas, illustrative studies, and the main open questions that recur across the mapped link patterns. The 'future research' column is intended as a compact, evidence-informed agenda derived from the stream structure rather than as definitive prescriptions. For DSS researchers, this stream summary can also serve as an organising lens for designing and evaluating AI-enabled decision support.

In the following discussion, we use this stream-based synthesis to position our contribution in relation to prior HAIC scholarship and to clarify the theoretical and practical implications of the review.

5. Discussion

This study harmonised the conceptual vocabulary of human – AI collaboration (HAIC) across Information Systems (IS) and Computer Science (CS) and linked these conceptualisations to how collaboration is evaluated in prior work. By consolidating definitional evidence where it is consistently available and by mapping cross-theme relatedness through co-occurrence and normalised overlap, the review offers an empirically grounded view of what structures the HAIC discourse and how evaluation logics travel with different conceptual framings.

Our synthesis suggests that despite a wide range of labels used to describe human – AI work only a relatively small subset of concepts emerges as theoretically and empirically central in the cross-disciplinary corpus. This work extends Vaccaro et al. (2024) and Berretta et al. (2023) by moving beyond conceptual discussions towards a systematic cross-disciplinary analysis, identifying the core constructs and boundary conditions that structure how human – AI collaboration is defined and differentiated.

Our quantitative mapping further demonstrates the extent to which these concepts are related, moving beyond narrative, conceptually fragmented accounts of the field (Baer et al., 2025) towards an evidence-grounded mapping of recurring link patterns. In particular, the co-occurrence structure shows that evaluation themes operate as a recurrent connector across multiple conceptual framings rather than as a peripheral methodological detail, and that some mechanism linkages remain selective even after normalising for prevalence (e.g. tight coupling between reliance dynamics and informational support; Kim et al., 2023). In this way, the study strengthens interpretability by showing which concerns reliably travel together across studies and which associations may be driven by broad theme popularity. The stream structure should also be read in light of corpus composition. Because the search targets studies that explicitly engage with HAIC conceptualisation or evaluation, CS work is more heavily represented and may exert greater influence on the co-occurrence patterns. At the same time, the stream candidates are not purely CS-specific. Each stream contains some IS studies (A: 6/42; B: 6/33; C: 2/39; D: 3/13, IS/CS). This pattern suggests that the streams capture recurring cross-disciplinary concerns, while the relative weight of evidence differs by stream and is particularly limited for trust-calibration work in IS outlets. Accordingly, we treat the streams as indicative organising lenses for the retrieved corpus rather than as balanced population estimates across IS and CS. Taken together, the corpus composition also suggests a gap in IS research. Within our retrieved corpus, comparatively fewer IS studies foreground HAIC conceptualisation and, in particular, evaluation as primary objects of analysis. This under-representation matters for cumulative progress in IS because construct clarity and assessment logic are central for comparing findings across organisational settings.

In DSS research, Table 1 helps distinguish collaboration constructs from collaboration modes, reducing the risk of treating one-way tool use or automation substitution as 'collaboration' in design and evaluation claims (Metcalfe et al., 2021). The identified research streams provide guidance for the design and evaluation of AI-based decision support systems by highlighting recurring aspects emphasised in the HAIC literature, such as trust calibration, responsibility boundaries, and evaluation goals. This guidance should be read as an evidence-informed organising lens derived from the retrieved corpus, rather than as a prescriptive design standard. Against this backdrop, the theoretical and practical implications of these findings are threefold. First, for the HAIC body of knowledge, the results support more cumulative progress by clarifying a shared vocabulary and boundary conditions that distinguish collaboration from one-way tool use and automation-oriented substitution, thereby reducing false comparability (Berretta et al., 2023) across studies that use different labels for non-equivalent arrangements. Second, for evaluation research, the mapped link patterns suggest that what counts as 'collaboration quality' is often implicitly set by measurement choices. Prior work cautions that when outcomes are reported without clarifying the assumed collaboration mode, evaluation becomes fragile and cross-study comparisons can be misleading (Fragiadakis et al., 2024). This makes concept-to-measure alignment especially important when comparing tasks that differ in uncertainty, time horizons, and accountability demands. Third, the stream structure provides an organising lens for synthesising how outcome-oriented goals, governance arrangements, reliance management, and interaction mechanisms are jointly discussed in the HAIC literature. To make this structure actionable, Table 3 consolidates each stream's key ideas, illustrative studies, and the recurring open questions suggested by the link patterns; this agenda is intended as evidence-informed guidance rather than definitive prescriptions.

6. Conclusion

This review synthesises how IS and CS research conceptualises HAIC and how those conceptualisations align with evaluation emphases. Drawing on 149 studies published from 2018 to 2025, we used a concept-centric synthesis to stabilise core terms and boundary conditions, and we mapped theme relatedness using a co-occurrence matrix with a normalised association indicator. The resulting picture suggests a two-level structure. Some concepts frame what collaboration is and what it is for, while others describe the

mechanisms and governance conditions through which collaboration unfolds and is assessed. These patterns, in turn, motivate four preliminary streams.

Several limitations shape how these findings should be read. Our corpus reflects the coverage of the selected databases and the reach of our search strings. As a result, IS outlets are less represented than CS outlets in the final sample, so the mapped link patterns should be interpreted as indicative of the retrieved corpus rather than as population estimates across the two domains. Co-occurrence captures association, not causality, so stream boundaries should be treated as an interpretive map rather than evidence that one theme drives another. The coding scheme also compresses nuance. A theme dictionary can merge distinct meanings under one label, and the Jaccard index can be sensitive when themes are rare. Although we used boundary statements and a sensitivity check to reduce hub-driven effects, these steps cannot eliminate all dependence on coding choices.

Future work can deepen stream-level interpretation through closer reading and exemplar selection, with particular attention to how concepts are applied and operationalised across different task settings, and where their meaning or measurement varies. Building on the concept table and the co-occurrence structure, we plan to develop a multi-criteria HAIC evaluation framework that links outcomes to process and governance dimensions and makes collaboration-mode assumptions explicit; these extensions are outside the scope of the present paper. This study contributes in two ways. First, it begins to harmonise HAIC concepts across IS and CS by establishing a shared vocabulary and specifying boundary conditions that distinguish collaboration from one-way tool use and from automation that replaces human work, offering a clearer basis for cumulative theorising and more consistent construct use across studies. Second, it provides a transparent mapping of how conceptual and evaluative themes cluster across IS and CS studies, offering a grounded starting point for stronger theory building and more comparable assessment designs.

Author contributions

CRediT: **Yu Wang:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing; **Danielly de Paula:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing; **Ciara Heavin:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Yu Wang http://orcid.org/0009-0002-3100-8331
Danielly de Paula http://orcid.org/0000-0002-4837-7347
Ciara Heavin http://orcid.org/0000-0001-8237-3350

References

AI Index Steering Committee. (2025). AI index report 2025. Stanford Institute for Human-Centered Artificial Intelligence (HAI). https://hai.stanford.edu/ai-index/2025-ai-index-report

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. Association for Computing Machinery. https://doi.org/10.1145/3290605.3300233

Baer, I., Waardenburg, L., & Huysman, M. (2025). What is augmented? A metanarrative review of AI-based augmentation. *Journal of the Association for Information Systems*, *26*(3), 760–798. https://doi.org/10.17705/1jais.00921

Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C., & Kluge, A. (2023). Defining human AI teaming the human centered way: A scoping review and network analysis. *Frontiers in Artificial Intelligence*, *6*, 1250725. https://doi.org/10.3389/frai.2023.1250725

Caldwell, S., Sweetser, P., O'Donnell, N., Knight, M.J., Aitchison, M., Gedeon, T., Johnson, D., Brereton, M., Gallagher, M., & Conroy, D. (2022). An agile new research framework for hybrid human AI teaming: Trust, transparency, and

transferability. *ACM Transactions on Interactive Intelligent Systems*, *12*(3), Article 17. 1–36. https://doi.org/10.1145/3514257

Chen, Y.-C., Liu, H., & Wang, Y.-F. (2024). Governance design of collaborative intelligence for public policy and services. In *Proceedings of the 25th Annual International Conference on Digital Government Research* (pp. 146–155). Association for Computing Machinery. https://doi.org/10.1145/3657054.3657075

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J.M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, *61*(5), 637–643. https://doi.org/10.1007/s12599-019-00595-2

Duan, W., Zhou, S., Scalia, M.J., Yin, X., Weng, N., Zhang, R., Freeman, G., McNeese, N., Gorman, J., & Tolston, M. (2024). Understanding the evolvement of trust over time within human-AI teams. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW2), 521. https://doi.org/10.1145/3687060

Dubey, A., Abhinav, K., Jain, S., Arora, V., & Puttaveerana, A. (2020). HACO: A framework for developing human-AI teaming. In *Proceedings of the 13th Innovations in Software Engineering Conference (Formerly known as India Software Engineering Conference) (ISEC '20)* (Article 10, p. 9). Association for Computing Machinery. https://doi.org/10.1145/3385032.3385044

Fabri, L., Häckel, B., Oberlaender, A.M., Rieg, M., & Stohr, A. (2023). Disentangling human-AI hybrids: Conceptualizing the interworking of humans and AI-enabled systems. *Business & Information Systems Engineering*, *65*(6), 623–641. https://doi.org/10.1007/s12599-023-00810-1

Fletcher, S., & Islam, M.Z. (2018). Comparing sets of patterns with the Jaccard index. *Australasian Journal of Information Systems*, 22. https://doi.org/10.3127/ajis.v22i0.1538

Fragiadakis, G., Diou, C., Kousiouris, G., & Nikolaidou, M. (2024). Evaluating human AI collaboration: A review and methodological framework [preprint]. arXiv. https://doi.org/10.48550/arXiv.2407.19098

Fuchs, A., Passarella, A., & Conti, M. (2024). Optimizing delegation in collaborative human-AI hybrid teams. *ACM Transactions on Autonomous and Adaptive Systems*, *19*(4), Article 23. https://doi.org/10.1145/3687130

Goktas, P. (2024). Ethics, transparency, and explainability in generative AI decision-making systems: A comprehensive bibliometric study. *Journal of Decision Systems*, 1–29. https://doi.org/10.1080/12460125.2024.2410042

Gomez, C., Cho, S.M., Ke, S., Huang, C.-M., & Unberath, M. (2025). Human AI collaboration is not very collaborative yet: A taxonomy of interaction patterns in AI assisted decision making from a systematic review. *Frontiers in Computer Science*, *6*, 1521066. https://doi.org/10.3389/fcomp.2024.1521066

He, Z., Song, Y., Zhou, S., & Cai, Z. (2023). Interaction of thoughts: Towards mediating task assignment in human AI cooperation with a capability aware shared mental model. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Article 353, pp. 1–18). Association for Computing Machinery. https://doi.org/10.1145/3544548.3580983

Hemmer, P., Westphal, M., Schemmer, M., Vetter, S., Vössing, M., & Satzger, G. (2023). Human AI collaboration: The effect of AI delegation on human task performance and task satisfaction. In *Proceedings of the 28th International Conference on Intelligent User Interfaces* (pp. 453–463). Association for Computing Machinery. https://doi.org/10.1145/3581641.3584052

Jaccard, P. (1901). Etude de la distribution florale dans une portion des Alpes et du Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, *37*, 547–579. https://doi.org/10.5169/seals-266450

Jarrahi, M.H. (2018). Artificial intelligence and the future of work: Human AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Jha, A., Gupta, H., & K, S.G. (2024). Aqi prediction using human-AI teaming. In *Roceedings of the 2024 International Conference on IoT, Communication and Automation Technology (ICICAT)* (pp. 260–265). IEEE. https://doi.org/10.1109/ICICAT62666.2024.10922978

Kim, S.S.Y., Watkins, E.A., Russakovsky, O., Fong, R., & Monroy-Hernández, A. (2023). "Help me help the AI": Understanding how explainability can support human-AI interaction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23, article 250)* (p. 17). Association for Computing Machinery. https://doi.org/10.1145/3544548.3581001

Krausman, A., Neubauer, C., Forster, D., Lakhmani, S., Baker, A.L., Fitzhugh, S.M., Gremillion, G., Wright, J.L., Metcalfe, J.S., & Schaefer, K.E. (2022). Trust measurement in human-autonomy teams: Development of a conceptual toolkit. *Journal of Human-Robot Interaction*, *11*(3), Article 33. 1–58. https://doi.org/10.1145/3530874

Krinkin, K., Shichkina, Y., & Ignatyev, A. (2021). Co-evolutionary hybrid intelligence. In *Proceedings of the 2021 5th Scientific School "Dynamics of Complex Networks and their Applications" (DCNA 2021)* (pp. 112–115). IEEE. https://doi.org/10.1109/DCNA53427.2021.9587002

Lai, V., Chen, C., Smith-Renner, A., Liao, Q.V., & Tan, C. (2023). Towards a science of human-AI decision making: An overview of design space in empirical human-subject studies. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1369–1385). Association for Computing Machinery. https://doi.org/10.1145/3593013.3594087

Li, J., Yang, Y., Liao, Q.V., Zhang, J., & Lee, Y.-C. (2025). As confidence aligns: Understanding the effect of AI confidence on human self-confidence in human-AI decision making. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (p. 1111). Association for Computing Machinery. https://doi.org/10.1145/3706598.3713336

Metcalfe, J.S., Perelman, B.S., Boothe, D.L., & McDowell, K. (2021). Systemic oversimplification limits the potential for human AI partnership. *IEEE Access*, *9*, 70242–70260. https://doi.org/10.1109/ACCESS.2021.3078298

Muller, M., & Weisz, J. (2022). Extending a human-AI collaboration framework with dynamism and sociality. In *Proceedings of the 1st Annual Meeting of the Symposium on Human-Computer Interaction for Work* (Article 10, p. 12). Association for Computing Machinery. https://doi.org/10.1145/3533406.3533407

Pareek, S., Velloso, E., & Goncalves, J. (2024). Trust development and repair in AI-assisted decision-making during complementary expertise. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 546–561). Association for Computing Machinery. https://doi.org/10.1145/3630106.3658924

Paul, S., Yuan, L., Jain, H.K., Robert, L.P., Spohrer, J., & Lifshitz-Assaf, H. (2022). Intelligence augmentation: Human factors in AI and future of work. *AIS Transactions on Human-Computer Interaction*, *14*(3), 426–445. https://doi.org/10.17705/1thci.00174

Pinski, M., Adam, M., & Benlian, A. (2023). AI knowledge: Improving AI delegation through human enablement. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Article 25, pp. 1–17). Association for Computing Machinery. https://doi.org/10.1145/3544548.3580794

Puerta-Beldarrain, M., Gómez-Carmona, O., Sánchez-Corcuera, R., Casado-Mansilla, D., López de Ipiña, D., & Chen, L. (2025). A multifaceted vision of the human-AI collaboration: A comprehensive review. *IEEE Access*, *13*, 29375–29405. https://doi.org/10.1109/ACCESS.2025.3536095

Rezwana, J., & Maher, M.L. (2023). Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems. *ACM Transactions on Computer-Human Interaction*, *30*(5), 67. https://doi.org/10.1145/3519026

Rockbach, J.D., Fuchs, S., Witte, T.E.F., & Bluhm, L.-F. (2022). Ingredients for hybrid intelligence: Towards an integrated theory and application. In *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICHMS56717.2022.9980541

Sadeghian, S., Uhde, A., & Hassenzahl, M. (2024). The soul of work: Evaluation of job meaningfulness and accountability in human-AI collaboration. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW1), 130. https://doi.org/10.1145/3637407

Saup, T.O., Asghar, J., Kanbach, D.K., & Kraus, S. (2026). From pilots to decision systems: Embedding generative AI into strategic decision-making through a socio-technical and governance lens. *Journal of Decision Systems*, *35*(1). https://doi.org/10.1080/12460125.2025.2597835

Sawyer, S., & Jarrahi, M.H. (2014). Sociotechnical approaches to the study of information systems. In H. Topi & A. Tucker (Eds.), *Computing handbook* (3rd ed. pp. 5–1–5–27). Chapman & Hall/CRC.

Schelble, B. G., Flathmann, C., Musick, G., McNeese, N. J., & Freeman, G. (2022). I see you: Examining the role of spatial information in human-agent teams. *Proceedings of the ACM on Human-Computer Interaction*, *6*(CSCW2), 374. https://doi.org/10.1145/3555099

Schlup, B., & Corradini, A. (2024). A flexible multi-agent systems task environment for simulating hybrid intelligence. In *Proceedings of the 2024 IEEE 12th International Conference on Intelligent Systems (IS)* (pp. 1–7). IEEE. https://doi.org/10.1109/IS61756.2024.10705173

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, *36*(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Spitzer, P., Holstein, J., Hemmer, P., Vössing, M., Kühl, N., Martin, D., & Satzger, G. (2025). Human delegation behavior in human-AI collaboration: The effect of contextual information. *Proceedings of the ACM on Human-Computer Interaction*, *9*(CSCW1), Article CSCW101. 1–28. https://doi.org/10.1145/3710999

Teodorescu, M., Morse, L., Awwad, Y., & Kane, G.C. (2021). Failures of fairness in automation require a deeper understanding of human-ML augmentation. *MIS Quarterly*, *45*(3), 1483–1500. https://doi.org/10.25300/MISQ/2021/16535

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, *8*(12), 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

Van der Aalst, W.M.P. (2021). Hybrid intelligence: To automate or not to automate, that is the question. *International Journal of Information Systems and Project Management*, *9*(2), 5–20. https://doi.org/10.12821/ijispm090201

Vereschak, O., Bailly, G., & Caramiaux, B. (2021). How to evaluate trust in AI-assisted decision making? A survey of empirical methodologies. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), Article 327, 1–39. https://doi.org/10.1145/3476068

Wang, B.Y., Boell, S.K., Riemer, K., & Peter, S. (2023). Human agency in AI configurations supporting organizational decision-making. In *Acis, 2023 Proceedings (Paper 53)*. AIS eLibrary. https://aisel.aisnet.org/acis2023/53

Webster, J., & Watson, R.T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), xiii–xxiii. https://www.jstor.org/stable/4132319

Wegner, P. (1976). Research paradigms in computer science. In *Proceedings of the 2nd International Conference on Software Engineering (ICSE '76)* (pp. 322–330). IEEE Computer Society Press.

Zabaleta, K., Lago, A. B., López De Ipiña, D., DiModica, G., Santos De La Cámara, R., & Pistore, M. (2019). Combining human and machine intelligence to foster wider adoption of e-services. In *2019 IEEE SmartWorld Conference* (pp. 1854–1859). IEEE. https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00326

Zhang, Y., Liao, Q. V., & Bellamy, R. K. E. (2020). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT '20)* (pp. 295–305). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372852

Zhou, L., Paul, S., Demirkan, H., Yuan, L., Spohrer, J., Zhou, M., & Basu, J. (2021). Intelligence augmentation: Towards building human-machine symbiotic relationship. *AIS Transactions on Human-Computer Interaction*, *13*(2), 243–264. https://doi.org/10.17705/1thci.00149

Zhou, L., Rudin, C., Gombolay, M., Spohrer, J., Zhou, M., & Paul, S. (2023). From artificial intelligence (AI) to intelligence augmentation (IA): Design principles, potential risks, and emerging issues. *AIS Transactions on Human-Computer Interaction*, *15*(1), 111–135. https://doi.org/10.17705/1thci.00185