



Article

# Trust in Algorithms in E-Commerce Recommender Systems: A Bibliometric Mapping (2012–2025) and a Managerial Playbook for Acceptability

Marija Gombar <sup>1</sup> , Amir Topalović <sup>2</sup> and Mirjana Pejić Bach <sup>3,4,5,\*</sup>

<sup>1</sup> Ministry of Defence, 10000 Zagreb, Croatia; magombar@unin.hr

<sup>2</sup> Aisma Srl, 20122 Milan, Italy; amir.topalovic@aisma.it

<sup>3</sup> Faculty of Economics and Business, University of Zagreb, 10000 Zagreb, Croatia

<sup>4</sup> Faculty of Commercial and Business Sciences, Lava 7, 3000 Celje, Slovenia

<sup>5</sup> IEDC Bled School of Management, Prešernova Street 33, 4260 Bled, Slovenia

\* Correspondence: mpejic@net.efzg.hr

## Abstract

With the increasing integration of artificial intelligence into e-commerce platforms, trust in algorithmic decision-making has become a critical issue. Recommender systems significantly shape consumer choices and influence visibility within digital marketplaces, yet remain largely opaque. This study aims to bridge the gap between algorithmic accuracy and perceived trustworthiness by conducting a bibliometric and topic modeling analysis of 163 peer-reviewed publications (2012–2025). Results indicate a paradigmatic shift from usability-focused approaches toward governance-aware frameworks encompassing fairness, explainability, and accountability. To capture this transformation, the Acceptance Triangle model is introduced, conceptualising algorithmic acceptability across three interdependent layers: trust calibration at the interface level, exposure fairness at the platform level, and accountability mechanisms at the institutional level. The model is further operationalised through the Trust UX Playbook—nine managerial design levers with associated key performance indicators—and a Composite Acceptability Score integrating accuracy, fairness, and complaint reduction. The findings suggest that trust alone may be insufficient for understanding long-term acceptability in e-commerce recommender systems. Instead, the alignment between user experience, market equity, and governance legitimacy is interpreted as an analytically useful condition for conceptualising algorithmic acceptability. This research contributes a structured framework for assessing and designing acceptable recommender systems, offering actionable guidance for designers, decision-makers, and regulatory stakeholders seeking to improve algorithmic transparency, fairness, and accountability in online commerce.

**Keywords:** algorithmic trust; digital governance; e-commerce; explainable AI; exposure fairness; recommender systems; trust calibration



Academic Editor: Albert Y.S. Lam

Received: 29 April 2026

Revised: 19 May 2026

Accepted: 22 May 2026

Published: 27 May 2026

**Copyright:** © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

E-commerce recommender systems have become central infrastructures of digital marketplaces. By ranking products, personalising offers, and shaping visibility, they influence not only consumer choice but also the economic opportunities of sellers and brands. Their value has traditionally been assessed through technical and behavioural indicators such as prediction accuracy, click-through rates, conversion, and retention.

However, these performance-oriented measures do not fully capture whether users perceive recommendations as transparent, fair, controllable, and legitimate. As recommender systems increasingly mediate commercial attention and platform visibility, the central problem is no longer only whether recommendations are accurate, but whether they are acceptable.

This problem is situated within a broader transformation in which artificial intelligence increasingly reshapes organisational processes, consumer environments, and decision-making systems. AI is no longer limited to task automation, but is embedded in hybrid forms of human–AI collaboration that affect strategic, managerial, and operational domains [1]. At the same time, automation restructures work, expertise, and discretion, producing both displacement and task transformation [2], while algorithmic control increasingly reorganises decision-making and oversight in data-driven environments [3]. These developments have intensified ethical and epistemological concerns because AI systems mediate what is recognised as relevant, credible, or actionable knowledge [4,5]. Scholarship on augmentation further shows that systems designed to support human performance may also increase dependence, weaken skills, or reinforce surveillance-oriented forms of control [6]. These concerns are amplified by generative and opaque AI systems, which raise risks related to trust, manipulation, accountability, and machine behaviour [7–9]. More broadly, algorithmic systems are not neutral intermediaries: they may reproduce structural bias, obscure responsibility, and reinforce unequal outcomes across social and institutional settings [10–15]. Such dynamics create conditions of epistemic asymmetry, in which users are affected by algorithmic decisions without being able to fully understand, contest, or recalibrate them [16].

In e-commerce recommender systems, these broader concerns become concrete and measurable. Attempts to foster trust through technical explainability often fail when they do not account for the social, commercial, and institutional conditions that mediate acceptance or rejection of algorithmic authority [17]. Algorithmic trust therefore cannot be reduced to a functional parameter, but should be understood as a dynamic negotiation between transparency, control, fairness, and perceived legitimacy [18,19]. AI-powered recommender systems do not merely reflect user preferences; they actively shape them, raising concerns about manipulation, autonomy, informed choice, and the fair distribution of visibility in digital marketplaces [20,21].

In contrast to earlier frameworks that conceptualised user trust as a function of perceived usefulness and ease of use [22], contemporary studies stress the need for a multi-dimensional understanding of algorithmic acceptability, encompassing fairness, explainability, transparency, and governance [23,24]. These dimensions are particularly relevant in high-stakes contexts where algorithmic recommendations influence financial decisions, health behaviours, or political views [25]. Moreover, a persistent gap exists between technical performance metrics and actual user trust. Systems optimised for accuracy or click-through rate often fail to address cognitive, emotional, or moral dimensions of user acceptance [26]. Bridging this gap requires not only technical improvements, but a conceptual reframing of what constitutes an acceptable algorithmic decision—from the perspective of both individuals and society [27].

To address the gap between algorithmic accuracy and algorithmic acceptability, this study introduces the Acceptance Triangle as its primary conceptual contribution. The model conceptualises acceptability in e-commerce recommender systems as the alignment of three interdependent layers: trust calibration at the user-interface level, exposure equity at the marketplace level, and accountability by design at the governance level. While existing literature often examines technical, ethical, or perceptual dimensions separately [21], the proposed model integrates them into a unified framework for e-commerce recommender

systems. It draws on complementary research streams, including human–AI interaction [1], fairness in algorithmic ranking [22], and the ethics of machine decision-making [23].

The study makes three contributions. First, it develops the Acceptance Triangle as a conceptual model of algorithmic acceptability in e-commerce recommender systems. Second, it uses bibliometric mapping and BERTopic topic modelling of 163 peer-reviewed publications from 2012 to 2025 to empirically ground the conceptual synthesis and identify the field's intellectual, thematic, and temporal structure. Third, it translates the model into operational guidance through the Trust UX Playbook and an exploratory Composite Acceptability Score (CAS), which are proposed as managerial instruments for future empirical testing rather than as fully validated universal metrics. In this sense, the article combines diagnostic mapping with conceptual synthesis and practical operationalisation, contributing to current debates on algorithmic governance, trustworthiness, and user agency [8,14,17].

Guided by these aims, the study addresses four research questions: (i) RQ1: How has research connecting trust, transparency, fairness, and explainability in e-commerce recommender systems evolved between 2012 and 2025?; (ii) RQ2: Which intellectual, thematic, and methodological clusters structure this field?; (iii) RQ3: How are user-level trust calibration, marketplace-level exposure equity, and governance-level accountability represented and connected in the bibliometric and topic-modelling results?; and (iv) RQ4: How can these patterns inform a conceptual framework and managerial playbook for algorithmic acceptability in e-commerce recommender systems?

The remainder of this article is structured as follows. Section 2 reviews the literature on algorithmic trust, transparency, fairness, accountability, and cognitive-perceptual factors in trust formation. Section 3 details the bibliometric and topic-modelling methodology. Section 4 reports the descriptive, relational, and semantic findings. Section 5 develops the Acceptance Triangle, explains how it is grounded in the mapped structures, and translates it into the Trust UX Playbook and the exploratory CAS. Section 6 concludes with limitations and directions for future research.

## 2. Literature Review

### 2.1. Trust in Algorithmic Decision-Making

Trust in algorithmic decision-making has emerged as a critical concern across disciplines, particularly as algorithmic systems are increasingly embedded in contexts that demand reliability, fairness, and accountability. Unlike traditional interpersonal trust, which is built through repeated social interactions, trust in algorithms operates at the intersection of perceived competence, epistemic authority, interface design, and regulatory assurance. Users often delegate decision-making to algorithms, based on assumptions of objectivity, consistency, and data-driven accuracy, even without a full understanding of how these systems function [28]. This delegation reflects a fundamental shift from interpersonal judgment to reliance on computation, often driven by the promise of efficiency and reduced human error. However, trust in algorithms is not merely instrumental; it also reflects an epistemic stance whereby users attribute superior knowledge or insight to algorithmic outputs, particularly in data-rich domains. The perception of algorithms as neutral or unbiased agents can lead to inflated expectations of accuracy and rationality [29]. Such trust, though, may be misplaced if it obscures the biases embedded in training data, model architectures, or optimization objectives. Indeed, trust in algorithms often rests on a paradox: users may place confidence in systems precisely because they lack insight into their inner workings. This opacity-induced confidence, while stabilizing user acceptance in the short term, can undermine critical engagement and long-term accountability.

Moral sensitivity further complicates trust in algorithmic systems. Users are more likely to reject algorithmic decisions they perceive as morally inappropriate, even when those decisions are technically correct or statistically optimal [30]. This divergence between computational rationality and moral intuitiveness points to a gap between algorithmic design logics and human ethical expectations. Moreover, trust is rarely static. It is subject to dynamic calibration based on experience, context, and perceived reliability. Empirical studies suggest that users adjust their level of trust in response to feedback, transparency cues, and perceived controllability of the system [31]. Systems that allow for user intervention, explanations, or corrections tend to foster calibrated trust, where reliance aligns with observed performance. Explainability plays a crucial role in shaping algorithmic trust, particularly for non-expert users. Interface design, information granularity, and linguistic framing influence how users interpret algorithmic decisions and whether they perceive the system as trustworthy [32]. Trust, therefore, is not only a function of algorithmic accuracy, but also of communicative fluency between the system and its users. Beyond individual perception, trust in algorithms is embedded within socio-cultural contexts. Public attitudes toward algorithmic decision-making vary significantly across regions, influenced by historical experiences with technology, cultural values, and governance structures [33]. Regulatory environments also frame trust by signaling which systems are deemed legitimate or safe.

The institutionalization of trust through policy further emphasizes its systemic role. Legal scholarship on patterned inequality and algorithmic prediction further shows that trust is no longer merely a subjective user attitude, but is increasingly connected to the distributive consequences of automated decision-making systems [34]. Regulatory mechanisms such as transparency mandates, risk classification, and auditing procedures aim to operationalize trustworthiness as a measurable and enforceable standard. This shift reframes trust from an emergent outcome to a design constraint, aligning legal, technical, and ethical domains. Algorithmic fairness is another axis along which trust is evaluated. Users tend to withdraw trust when they perceive outputs as discriminatory or systematically biased, regardless of the system's technical efficacy [35]. Fairness, therefore, is not an optional enhancement but a prerequisite for sustained trust. Scholars have argued that fairness must be embedded at the level of model design and data selection, not just at the level of interpretability or post hoc explanation [36]. In this view, trust becomes inseparable from justice, reinforcing the notion that algorithmic systems are not only technical instruments but also political artifacts with distributive consequences.

Beyond fairness and regulation, trust is also shaped by emotional and affective responses. Affective responses such as anxiety, uncertainty, or reduced emotional reassurance can diminish trust in AI systems even when users recognize their functional value [37]. Emotional resonance—the sense that a system responds to or aligns with user values and expectations—emerges as a critical dimension of trust formation. This affective layer is particularly relevant in human–AI interaction scenarios where the stakes are personal, such as healthcare recommendations or educational guidance. As such, trust is simultaneously rational, ethical, cultural, and emotional—a multidimensional construct that resists reduction to technical performance metrics.

Finally, trust in algorithms must be understood as a composite of systemic design, user perception, and contextual modulation. It involves not only what the algorithm does, but how it explains, adapts, and positions itself in relation to human values and institutional norms. In practical terms, fostering trust in algorithmic systems demands more than improving accuracy; it requires embedding accountability, transparency, fairness, and emotional intelligence into the very architecture of decision-making technologies. Without

such an integrated approach, trust risks becoming either superficial or brittle—insufficient to support long-term acceptance or ethical robustness in algorithmic governance.

## 2.2. Algorithmic Transparency, Fairness, and Accountability

Algorithmic transparency, fairness, and accountability represent the normative and technical backbone of trustworthy AI systems. As algorithmic decisions permeate domains such as healthcare, employment, and finance, calls for transparency have intensified. Transparency is often invoked as a remedy for the black-box nature of many AI systems, with the assumption that visibility into system logic can foster trust, detect bias, and support regulatory oversight [38]. However, scholars warn that transparency alone does not guarantee accountability, particularly when explanations are inaccessible to non-expert users or framed in a way that obscures power asymmetries [39].

Technical advances in explainable AI (XAI) aim to make algorithmic outputs more understandable, yet surveys of black-box explanation methods also show that interpretability remains dependent on the model, explanation technique, and decision context [40]. In high-stakes settings, post-hoc explanations may be insufficient, particularly where an inherently interpretable model could provide a more reliable basis for scrutiny and responsible decision-making [41]. Fairness in ranking-based systems introduces a further challenge: because ranking positions distribute visibility and attention, recommender systems can generate unequal exposure for ranked items or providers over time [42]. User acceptance of algorithmic decisions is also context-dependent, as algorithm aversion tends to increase when tasks are perceived as more subjective or personally consequential [43]. In addition, public perceptions of algorithmic fairness do not necessarily align with a single technical definition of fairness; experimental evidence shows that individuals evaluate competing fairness definitions differently depending on the decision context and the information made salient [44]. More broadly, research on dark patterns demonstrates that interface design can steer user behaviour in ways that compromise meaningful agency, even when formal choices remain available [45]. Institutional responses to these challenges include the European Union's AI Act, which establishes a risk-based regulatory framework and introduces transparency, human-oversight, and accountability obligations for systems falling within regulated categories [46]. While the Act represents a major regulatory milestone, its implications for particular e-commerce recommender-system configurations still require domain-specific operationalisation.

The emergence of independent auditing bodies, algorithmic impact assessments, and fairness-by-design guidelines reflects attempts to operationalize abstract principles into actionable practices [47]. Yet, critics argue that these frameworks risk becoming technocratic rituals unless embedded within broader social and institutional change [48].

The complexity of algorithmic accountability lies in the entanglement of design choices, institutional settings, and epistemic authority. Systems trained on biased or incomplete data perpetuate historical inequities, regardless of technical accuracy. Addressing this requires shifting the focus from algorithmic outputs to the upstream decisions about what is measured, prioritized, and optimized. A relational approach to fairness recognizes that algorithms are embedded in socio-technical assemblages, where power, identity, and access are continually negotiated [49]. Beyond institutional mechanisms, user agency plays a critical role in sustaining algorithmic accountability. Studies show that explainability increases perceived control, especially when users are allowed to contest or override decisions [50]. However, in many systems, user participation remains superficial or tokenistic. True accountability demands the capacity for meaningful contestation, supported by clear procedural norms and accessible information architectures. Without such capacities, transparency becomes a hollow virtue and fairness a performative ideal.

The ethical implications of algorithmic systems cannot be decoupled from their infrastructural and economic contexts. Corporate incentives, data asymmetries, and geopolitical pressures shape what is built, how it is deployed, and whom it benefits. Thus, fairness and transparency are not only design goals but also political commitments that require institutional will and ethical vigilance [51]. The path to accountable AI is not linear; it requires ongoing negotiation between technical possibility, social legitimacy, and normative clarity. Embedding this triad into algorithmic development cycles is essential for mitigating harm and upholding justice in automated decision-making [52].

### 2.3. Ethical and Regulatory Considerations

Trust in AI-driven systems is increasingly scrutinised not only as a technical or behavioural phenomenon but as a normative and regulatory concern. The shift from usability-centered transparency to governance-centered explainability has positioned trust within a broader matrix of fairness, accountability, and legality.

In algorithmically mediated environments, trust extends beyond individual perception to encompass normative structures that govern transparency, fairness, and accountability. As algorithmic systems assume increasingly consequential roles in society—from financial decision-making to e-commerce curation—ethical and regulatory scrutiny has intensified. Scholars and policymakers alike have turned attention to the limitations of usability-driven explainability and advocated for normative frameworks that treat explainability as a governance instrument rather than a design feature [53,54]. The EU Artificial Intelligence Act operationalizes a risk-based taxonomy that mandates transparency, robustness, risk management, and accountability obligations for high-risk systems [46]. Complementing this regulatory context, cybersecurity policy work on critical sectors further underlines that accountability depends on risk-based governance, resilience obligations, and traceable oversight mechanisms [55].

In parallel, the General Data Protection Regulation (GDPR) and the Digital Services Act (DSA) establish a legal substrate for individual data rights, content moderation, and algorithmic accountability. This ecosystem of norms introduces tension between technical explainability and legal interpretability, with research highlighting gaps between what systems can explain and what individuals need to understand [56].

Transparency, as a regulatory objective, implies both epistemic accessibility and procedural auditability. However, explainability mechanisms often remain post hoc rationalizations that do not meet legal standards for meaningful information [57]. Scholars argue that the overreliance on black-box models undermines fairness, particularly in domains such as hiring, lending, or insurance [58,59]. Calls for fairness by design have emerged to shift the burden upstream—embedding normative constraints into model training and evaluation pipelines [60]. These include metrics for exposure equity, distributional parity, and representational harm, which challenge the traditional accuracy–efficiency trade-off at the core of machine learning. From a governance standpoint, platform self-regulation has proven insufficient to address structural biases. The lack of standardised audit protocols and the opaqueness of proprietary recommender systems impede oversight, even as platforms increasingly shape public discourse and consumption [61]. Academic proposals for algorithmic impact assessments draw inspiration from environmental and data protection precedents, advocating for anticipatory evaluation frameworks that integrate social, economic, and ethical consequences [62].

Ethical critiques further underscore the asymmetries of power inherent in data-driven persuasion. Trust-enhancing digital environments can also generate new vulnerabilities when perceived community or platform trust encourages users to disclose personal information more readily [63]. Related evidence indicates that anthropomorphic design and the

form of explanation provided by a biased AI system can shape whether users recognise bias and whether they continue to trust the system [64]. Beyond commercial applications, domain-specific research on explainable AI in cancer detection similarly demonstrates that transparency and interpretability must be adapted to the stakes and requirements of a particular decision context [65]. More broadly, research on privacy behaviour shows that users' decisions are shaped by uncertainty, context, and the susceptibility of privacy preferences to commercial or institutional influence [66]. Taken together, these findings reinforce the need to evaluate algorithmic acceptability not only through technical performance, but also through interface effects, exposure consequences, contextual vulnerability, and governance obligations.

As summarised in Table 1, prior work provides complementary foundations for understanding algorithmic acceptability, ranging from AI trust and recommender explainability to exposure fairness and regulatory accountability. However, these streams remain insufficiently integrated in research on e-commerce recommender systems. This gap motivates the present bibliometric approach, which maps how technical, behavioural, market-level, and legal dimensions converge within the study domain.

**Table 1.** Representative conceptual, empirical, and regulatory foundations relevant to algorithmic acceptability in e-commerce recommender systems.

Reference	Focus of the Study	Methodology	Relevance and Remaining Gap for This Study
[19]	Explainable recommendation systems, including transparency, trustworthiness, and explanation design.	Survey and conceptual synthesis.	Provides a recommender-system foundation, but does not map the convergence of trust, fairness, and regulation in e-commerce.
[29]	Determinants of human trust in artificial intelligence across empirical studies.	Review of empirical research.	Establishes the broader AI-trust foundation, but is not specific to e-commerce recommender systems.
[32]	Effects of e-commerce recommendation-system transparency on consumer trust.	Empirical modelling.	Directly addresses consumer trust in e-commerce recommendations, but remains focused on the user level rather than market-level fairness and governance.
[40]	Methods for explaining black-box models.	Survey and taxonomy of explanation methods.	Provides the XAI foundation, but does not address retail-platform trust or merchant visibility outcomes.
[41]	Use of interpretable models instead of post-hoc explanations in high-stakes decisions.	Perspective and evidence-based argument.	Strengthens the interpretability argument, but is not tailored to e-commerce recommender systems.
[42]	Equity of attention and amortised individual fairness in rankings.	Formal model and empirical evaluation.	Provides the exposure-fairness foundation, but does not directly connect ranking fairness with consumer trust.
[46]	Risk-based AI obligations concerning transparency, accountability, and oversight.	Legal and regulatory framework.	Provides the regulatory basis for accountability, but requires domain-specific operationalisation for recommender systems.
[53]	Automated decision-making under EU regulation and the right to explanation.	Legal analysis.	Establishes the legal relevance of explanation, but is not specifically embedded in e-commerce recommender-system design.

Abbreviations: To list representative studies only; to summarise focus, method, and gap in telegraphic form. To define abbreviations: RS = recommender systems; XAI = explainable AI. Table 1 provides a selective conceptual synthesis of representative studies already cited in the main text and included in the reference list.

Table 1 highlights how interface-level explanation works, linking to market-level exposure equity and governance-level accountability, motivating our cross-layer model of acceptability. As summarised in Table 1, prior studies have produced fragmented yet complementary insights—ranging from conceptual definitions of trust to empirical evaluations of explainability and fairness—yet no integrative mapping delineates how these domains converge within AI-driven e-commerce. Accordingly, this study undertakes a bibliometric analysis of trust and transparency in e-commerce recommendations. As Table 1 illustrates, trust, transparency, and fairness are deeply interwoven but insufficiently integrated across regulatory, technical, and behavioural domains. Having situated trust within ethical and governance discourses, the next section examines how users cognitively and emotionally engage with algorithmic systems in real-time interaction.

#### 2.4. Cognitive and Perceptual Factors in Trust Formation

Trust in algorithmic systems is not merely a rational evaluation but an emergent state shaped by cognitive heuristics, emotional cues, and perceived agency. Unlike interpersonal trust, which builds on observed behavior and social context, trust in AI is often anchored in system interfaces, symbolic signals, and user expectations formed under opacity and automation. Classical frameworks of trust—including ability, integrity, and benevolence—must be recalibrated for human–AI interaction, where explainability, perceived control, and reliability substitute for social cues [67]. In recommender systems, interface features and output framing significantly influence user acceptance, with studies showing that users prefer explanations that match their decision-making style [68,69]. Cognitive processing under algorithmic mediation is often constrained by a limited understanding of technical processes, giving rise to “automation bias” and misplaced reliance [70]. Trust propensity—the general tendency to trust technologies—interacts with perceptions of system usefulness and ease of use, especially in contexts of repeated exposure [71,72]. However, exposure itself is not neutral; algorithmic visibility, nudging, and personalization shape what users believe to be organic choices, reinforcing the illusion of autonomy [73]. Affective responses such as comfort, familiarity, or aesthetic satisfaction are routinely elicited through design, contributing to what has been described as algorithmic intimacy or simulated care [74,75].

The concept of algorithm aversion, where users reject systems after minor failures, highlights the volatility of trust in AI compared to human judgment [76,77]. Conversely, algorithm appreciation can emerge when the system is perceived as more objective or consistent than humans, especially in domains like medicine or finance [78,79]. Yet this appreciation often collapses when users sense a lack of transparency or accountability—particularly when black-box systems produce outcomes with high personal impact [80]. Perceptual trust is also modulated by the degree of control users feel they have. Research shows that systems offering optionality, feedback mechanisms, or recourse pathways are more likely to be trusted, even if underlying performance remains constant [81]. This has led to calls for participatory design models that foreground user agency, especially in contexts like e-commerce and employment where algorithmic outputs directly affect personal well-being [82,83]. At the same time, studies warn against conflating usability with trustworthiness; the ease of interaction can mask structural biases or hidden manipulation [84,85].

Digital literacy and prior experience play a significant role in shaping trust calibration. Users with higher technical literacy are better able to detect anomalies, question outputs, and resist overreliance, while novice users are more prone to cognitive shortcuts and passive acceptance [86]. Trust formation, therefore, emerges at the intersection of psychological predispositions, interface affordances, and system transparency—requiring multidimensional strategies for ethical design and governance.

### 3. Materials and Methods

To ensure transparency, replicability, and methodological rigor, this study follows the SPAR-4-SLR protocol, which is structured into four interdependent stages: assembly, arrangement, assessment, and reporting [87]. The approach is grounded in classic bibliometric foundations [88,89] and aligned with science mapping workflows that integrate performance metrics with network analytics [90,91]. To address potential methodological blind spots, multiple techniques are integrated—co-citation, bibliographic coupling, BERTopic, and lexicon-aligned overlays—combined with stability diagnostics such as ARI metrics. The post-2017 governance turn is explicitly tested using structured cluster interpretation. The methodological protocol and the corresponding selection flow are summarised in Tables 2 and 3, respectively.

**Table 2.** SPAR-4-SLR sampling and screening decisions (robust protocol view). Source: Scopus (Elsevier). Corpus retrieved 20 October 2025; authors’ calculations.

SPAR-4-SLR Stage	Decision	Implementation in this Study	Rationale	Evidence/Artifact (During Review)
Assembling	Index	Scopus (TITLE–ABS–KEY)	Consistent metadata for authors/affiliations/keywords in our domain	Scopus CSV/RIS export
	Query & date	(trust OR transparency OR fairness OR explainability) AND (“recommender system” OR “recommendation system”) AND (“e-commerce” OR “online retail”); 20 October 2025	Captures trust/XAI/fairness in retail RS	Query string + timestamp
Arranging	Language & type	English; Article, Conference Paper; 2012–2025	Field coverage; method-bearing items	Export with filters noted
	Domain filter	Keep e-commerce recommenders; remove off-domain (medical/education), editorials/notes, duplicates.	Align the corpus to retail trust/transparency	PRISMA-style counts
	Thesaurus	trust/trustworthiness; transparency/explainability/interpretability; e-commerce/online retail	Harmonize near-synonyms; reduce leakage	Thesaurus file
Assessing	Normalization & counting	Association-strength; full counting	Standard in VOSviewer; comparability	VOSviewer project files
	Clustering	Smart local moving	Stable community detection	Cluster config in project file
	Topic model	BERTopic (MiniLM; UMAP; HDBSCAN; c-TF-IDF)	Semantic drift & themes	Python notebook/params
Reporting	Stability & sensitivity	ARI ≥ 0.80 (co-citation), ≥0.76 (keywords); threshold sweeps	Guard against parameter artifacts	Stability logs/plots

**Table 3.** Selection flow and counts (PRISMA-style summary).

Step	Filter/Criterion	Description	Output Count
1	Title–Abstract–Keywords	TITLE-ABS-KEY. (trust OR transparency OR fairness OR explainability) AND (“recommender system” OR “recommendation system”) AND (“e-commerce” OR “online retail”); Years: 2012–2025	178
2	Language	Language: English	174
3	Document type	Document types: Article, Conference Paper	166
4	Domain relevance & duplicates	Remove off-domain (e.g., medical/education RS), editorials/notes; deduplicate	163

Abbreviations: Final corpus ( $n = 163$ ) used for all descriptive and network analyses. Source: Scopus (Elsevier); authors’ calculations.

The review domain is defined as AI-driven trust and transparency in e-commerce recommender systems. A two-phase workflow is implemented: (1) dataset extraction and cleaning, and (2) bibliometric and network analysis. Scopus is selected as the exclusive source index to ensure metadata consistency, especially regarding author affiliations, keywords, and funding declarations. To mitigate single-index bias, three validation mechanisms were conducted: (i) a 10% stratified manual screening of the corpus, (ii) a coverage audit of key journals over 2012–2025, and (iii) a thesaurus leakage test to assess incremental recall when applying near-synonyms. These checks did not indicate major coverage distortions in relation to the study objectives; however, they cannot eliminate the inherent limitations of relying on a single bibliographic database. Full TITLE-ABS-KEY query strings, export timestamps, filtering criteria, and the thesaurus file are provided to facilitate replicability.

To ensure conceptual coherence and methodological relevance, inclusion criteria were applied to limit the dataset to peer-reviewed journal articles and conference papers published in English between 2012 and 2025. This period captures the emergence and consolidation of key constructs such as explainability, algorithmic governance, and fairness of exposure in the domain of e-commerce recommender systems. Editorials, short notes, purely conceptual contributions lacking methodological grounding, duplicates, and off-domain documents (e.g., studies focused on medical or educational recommenders) were excluded. The initial query retrieved 178 records, of which 163 were retained after applying language and document-type filters, as well as domain-specific screening.

The final corpus size ( $n = 163$ ) is appropriate for the purpose of this study because the analysis is designed as a focused bibliometric and semantic mapping of a clearly delimited domain rather than as a population-level statistical inference. In BERTopic-based modelling, the aim was not to estimate universal topic prevalence, but to identify interpretable semantic structures within a thematically coherent corpus. The robustness of the topic solution was supported through parameter sensitivity checks, topic merging based on cosine similarity, exclusion of topics representing less than 1% of the corpus, positive NPMI coherence values for retained topics, and five bootstrap resamples. Accordingly, BERTopic is used as a complementary interpretive layer alongside co-citation, bibliographic coupling, and keyword co-occurrence analyses.

To ensure consistency and traceability, records were standardized across affiliations, institutions, countries, and author keywords through rule-based string normalization. Ambiguous cases were manually verified. Unique identifiers (DOIs and Scopus EIDs) were preserved throughout all preprocessing and analytical stages. Institutional and geographic metadata were harmonized using ISO-3 country codes and validated against GRID/ROR databases when available. Author names were disambiguated via initial-matching and ORCID cross-referencing. Keywords were normalized through case-folding, lemmatization, and singularization, and multiword expressions were unified according to a custom-built thesaurus. The full selection process and document counts are summarized in a PRISMA-style overview in Table 3.

To assess the comprehensiveness of the corpus, an internal coverage check was conducted. This included query variation using synonyms, lemmatised phrases, and alternative expressions; manual screening of borderline cases; and spot-checks of key authors and sources. The results did not indicate major coverage gaps in relation to the study objectives, although database-specific indexing constraints and the exclusion of non-English records remain relevant limitations. All metadata normalization steps—covering authors, affiliations, and keywords—were documented and logged to ensure full traceability. Terminology was aligned with governance-related concepts such as transparency, explainability, and fairness, as defined in established frameworks [92]. The thesaurus also incorporated

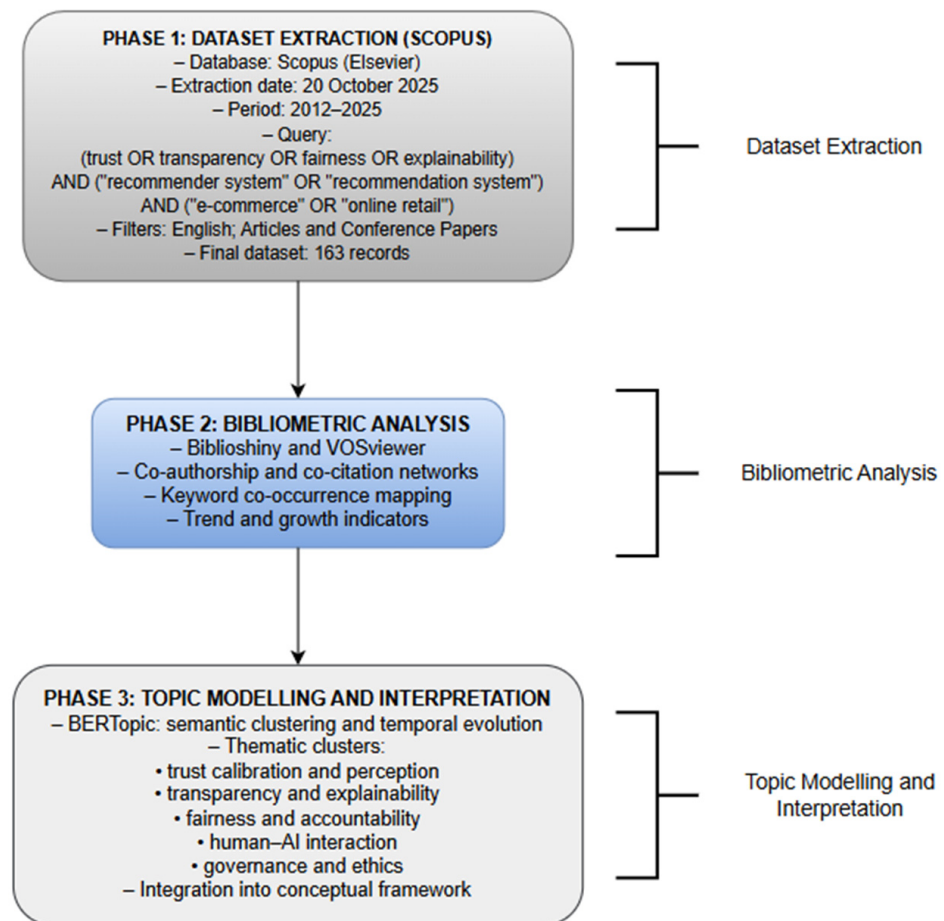
governance-adjacent UX terminology, including design manipulations such as “dark patterns”. Early notions of algorithmic transparency were adapted from foundational work in the recommender literature.

Descriptive indicators and conceptual structures were generated using Bibliometrix/Biblioshiny in R, while relational mapping was performed in VOSviewer. The descriptive analysis included annual publication trends, sources, authorship, geographic distribution, citation metrics, and normalized impact (e.g., citations per year). The structural analysis covered three dimensions: co-citation (intellectual foundations), bibliographic coupling (thematic proximity of recent publications), and keyword co-occurrence (topical evolution). Full counting and association-strength normalization were applied unless otherwise specified.

The analyses were conducted using R (version 4.3) with Bibliometrix/Biblioshiny (version 4.2), VOSviewer (version 1.6.2), and Python (version 3.11) with BERTopic (version 0.16). Topic modelling employed the sentence-transformer model all-MiniLM-L6-v2, UMAP, and HDBSCAN, with the analytical parameters reported below. As a robustness check, fractional counting was applied in parallel; the resulting clusters and key themes remained substantively unchanged. Stability was further confirmed by varying keyword frequency thresholds and comparing outputs from full and fractional counting approaches. No material thematic shifts were observed. This section follows the SPAR-4-SLR reporting convention, with transparent documentation of all filters, normalization steps, software configurations, and robustness diagnostics.

Acknowledged limitations include potential indexing updates after 20 October 2025; reliance on a single database (Scopus) with known coverage constraints; and the dynamic nature of topic- and time-dependent citation patterns, which may affect longitudinal comparability. This section is supported by Figure 1 and Tables 2 and 3. All query strings, filtering criteria, and software settings are reported in Section 3. Additional materials are available upon request. For community detection, the smart local moving (SLM) algorithm in VOSviewer was employed. Robustness was examined through variation of resolution parameters and edge thresholds, with cluster stability validated using density overlays and visual inspection. The resulting clusters were contrasted with known milestones in the field to support face validity. Resolution parameters were varied in the range of 0.5 to 1.2, with Adjusted Rand Index (ARI) values observed at  $\geq 0.80$  for co-citation and  $\geq 0.76$  for keyword networks, indicating high structural consistency.

An agenda-mapping approach was applied to distinguish early-stage clusters from more recent ones, allowing for the identification of thematic shifts, changes in institutional leadership, and the emergence of underrepresented topics. Keyword harmonization was guided by recent reviews on trustworthy recommender systems, including contemporary governance and explainability vocabularies. Topic modeling was performed using BERTopic, based on the sentence-transformer model all-MiniLM-L6-v2, UMAP (with  $n\_neighbors = 15$  and  $min\_dist = 0.0$ ), and HDBSCAN ( $min\_cluster\_size = 15$ ), with class-based TF-IDF used to extract representative terms. Topics covering less than 1% of the corpus were excluded, and highly similar topics (cosine similarity  $\geq 0.90$ ) were merged. Coherence was assessed using normalized pointwise mutual information (NPMI), with all retained topics achieving positive values. Topic stability was confirmed through five bootstrap resamples, with no topic falling below the retention threshold.



**Figure 1.** Study workflow (dataset extraction, bibliometric analysis, and topic modelling). Source: Scopus (Elsevier); authors' illustration and calculations.

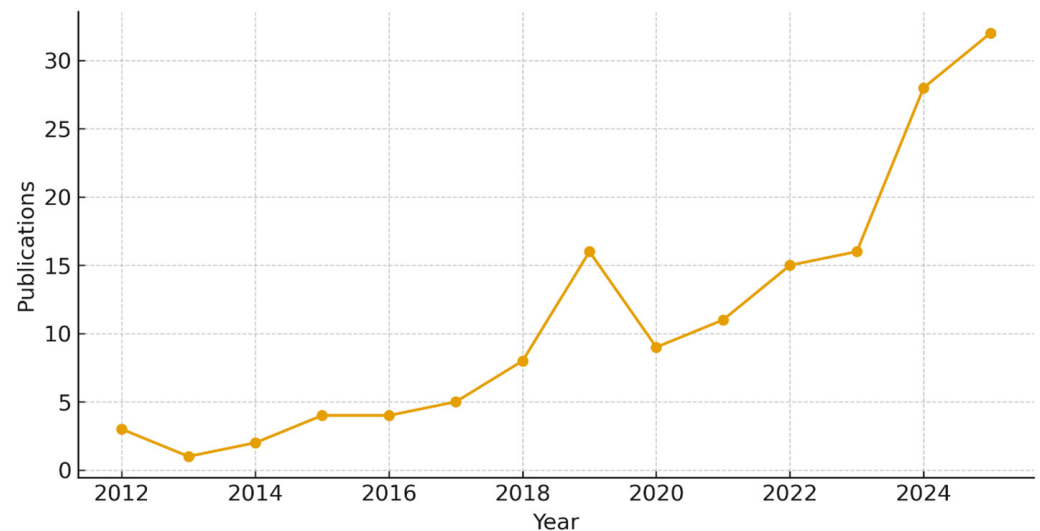
The topic-modelling output is used descriptively to support the interpretation of semantic shifts and thematic consolidation, rather than as an independent confirmatory test.

In addition to standard performance and relational analyses, three complementary analytical lenses were integrated: co-citation (intellectual foundations), bibliographic coupling (thematic proximity among recent works), and BERTopic (semantic drift and topic consolidation). A governance- and XAI-aligned lexicon was projected onto keyword co-occurrence networks to identify the post-2017 shift from explanation-as-usability toward explanation-as-governance. Robustness checks confirmed stable structures across parameter variations ( $ARI \geq 0.80$  for co-citation;  $\geq 0.76$  for author keywords).

## 4. Results

### 4.1. Descriptive Landscape (2012–2025)

The corpus (2012–2025;  $n = 163$ ) transitions from an HCI-centric niche to a governance-aware agenda. Output accelerates sharply after 2017, peaking in 2025 with 32 items; the inflection coincides with the consolidation of explainability and fairness debates in recommender research and with post-2018 uptake of accountability/compliance framings in retail platforms. Figure 2 reports year-by-year counts for the final Scopus corpus ( $n = 163$ ; retrieved 20 October 2025).



**Figure 2.** Annual publications in the study corpus, 2012–2025 (Scopus; retrieved 20 October 2025;  $n = 163$ ). Source: Scopus (Elsevier); authors' calculations.

The venue mix underscores this evolution: journals account for 61% (99/163) and conference papers for 39% (64/163). Lecture Notes in Computer Science anchors method-building contributions (5), while Technological Forecasting and Social Change and the International Journal of Human–Computer Interaction (3 each) reflect managerial and user-experience consolidation; flagship community venues (CHI, HICSS, 3 each) mark points where methodological advances meet applied retail contexts. Geographically, production concentrates in the United States, China, and India, followed by the United Kingdom and Germany, mapping a triad typical for information/HCI fields but here redirected toward trust calibration, exposure equity, and accountability in commerce.

Impact indicators mirror the shift from usability to acceptability. To reduce cohort bias, citation counts are reported as totals and citations-per-year, and dispersion is summarized with median and IQR; thresholds ( $\geq 10$ ;  $\geq 20$  citations) provide a coarse high-impact slice without over-weighting early cohorts. The set accumulates 3025 citations (median 4 per item; IQR 1–20) with a median  $\approx 1.33$  citations per year, consistent with citation latency for recent cohorts and with concentrated influence in papers published after 2019. Nearly two-fifths of items ( $\approx 39.3\%$ ) have  $\geq 10$  citations, and one-quarter ( $\approx 25.8\%$ ) have  $\geq 20$ , indicating a thickening mid-impact stratum alongside a small number of high-leverage works that connect explanation design to fairness, visibility, and compliance in ranking and exposure.

Annual growth patterns underpin this shift. After a long, low-volume phase, outputs rise steadily from 2017 onward and culminate in 2025 (32 items). To anchor the trajectory and avoid impressionistic claims, Figure 2 reports year-by-year counts for the final Scopus corpus ( $n = 163$ ; retrieved 20 October 2025), which we use as the baseline for subsequent dispersion and impact analyses. This temporal pattern is further contextualized by Table 4D, which outlines annual citation structures and confirms the citation latency effect for recent years.

The post-2017 inflection marks consolidation: explanation and fairness work moves from isolated prototypes to governance-aware studies, motivating a distributional impact view in Figure 3. Accordingly, Figure 3 profiles citation dispersion by year, showing how medians, IQRs, and tails evolve as the field expands.

**Table 4.** (A) Top sources by documents in the corpus (2012–2025;  $n = 163$ ). (B) Top authors by appearances in the Authors field (share of  $n = 163$ ). (C) Top affiliation countries by mentions in the Authors’ affiliation strings (share of  $n = 163$ ). (D) Annual citation structure of the corpus (2012–2025).

(A)			
Rank	Source Title	Docs	Share (%)
1	Lecture Notes in Computer Science	5	3.1
2	Conference on Human Factors in Computing Systems–Proceedings	3	1.8
3	Proceedings of the Annual HICSS	3	1.8
4	Technological Forecasting and Social Change	3	1.8
5	International Journal of Human–Computer Interaction	3	1.8
6	Studies in Science of Science	2	1.2
7	Wireless Personal Communications	2	1.2
8	Proceedings of SPIE–The Int’l Society for Optical Engineering	2	1.2
9	Advanced Materials Research	2	1.2
10	Procedia Computer Science	2	1.2

(B)			
Rank	Author	Docs	Share (%)
1	Shin, D.	6	3.7
2	Li, S.	3	1.8
3	Zhang, X.	3	1.8
4	Zhang, Z.	3	1.8
5	Swan, M.	3	1.8
6	Several authors at 2 appearances each	2	1.2

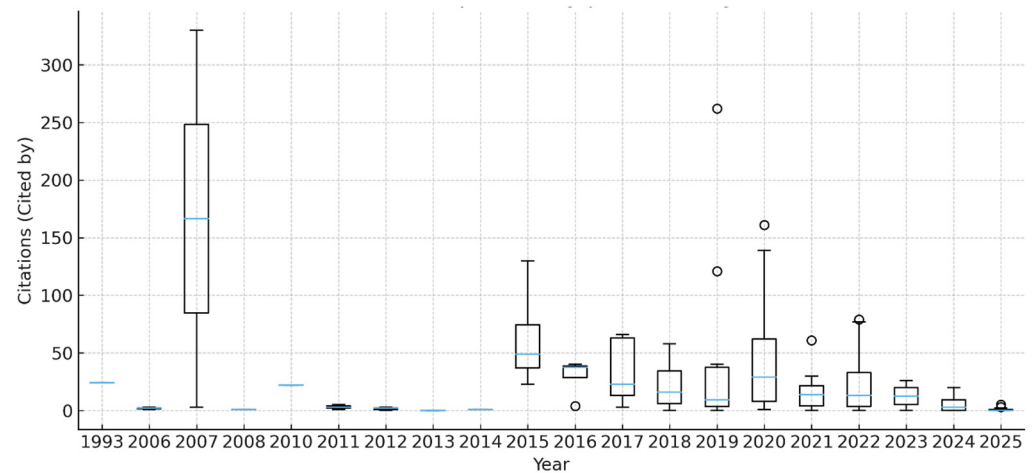
  

(C)			
Rank	Country	Mentions	Share (%)
1	United States	76	19.7
2	China	66	17.1
3	India	37	9.6
4	United Kingdom	29	7.5
5	Germany	22	5.7
6	Canada	16	4.2
7	Australia	14	3.6
8	United Arab Emirates	11	2.9
9	Netherlands	10	2.6
10	South Korea	9	2.3

(D)					
Year	TP	TC	TC/TP	% $\geq 10$	% $\geq 20$
2012	3	5	1.67	0.0	0.0
2013	1	0	0.00	0.0	0.0
2014	2	2	1.00	0.0	0.0
2015	4	251	62.75	100.0	100.0
2016	4	119	29.75	75.0	75.0
2017	5	39	7.80	20.0	0.0
2018	8	67	8.38	37.5	25.0
2019	16	302	18.88	68.8	37.5
2020	9	135	15.00	55.6	33.3
2021	11	189	17.18	63.6	36.4
2022	15	336	22.40	53.3	33.3
2023	16	199	12.44	56.3	31.3
2024	28	150	5.36	25.0	3.6
2025	32	21	0.66	0.0	0.0

Abbreviations: LNCS = Lecture Notes in Computer Science; TFSC = Technological Forecasting and Social Change; IJHCI = International Journal of Human–Computer Interaction. TP = total papers; TC = total citations; TC/TP = average citations per paper; %  $\geq 10$  = share of papers with  $\geq 10$  citations; %  $\geq 20$  = share with  $\geq 20$  citations.



**Figure 3.** Citation dispersion by publication year (Scopus; retrieved 20 October 2025). Boxes show the interquartile range (IQR), centre lines the median, whiskers denote Tukey fences, and dots indicate outliers. Source: Scopus (Elsevier); authors' calculations.

Figure 3 interprets Citation dispersion by publication year. Boxes show the interquartile range (IQR), center lines the median, and whiskers Tukey fences; dots indicate outliers. Dispersion thickens after 2019, consistent with citation latency for recent cohorts. Beyond widening IQRs, the corpus totals (3025 citations; median = 4; IQR = 1–20) and threshold shares ( $\geq 10 = 39.3\%$ ;  $\geq 20 = 25.8\%$ ) confirm a thickening mid-impact stratum after 2019.

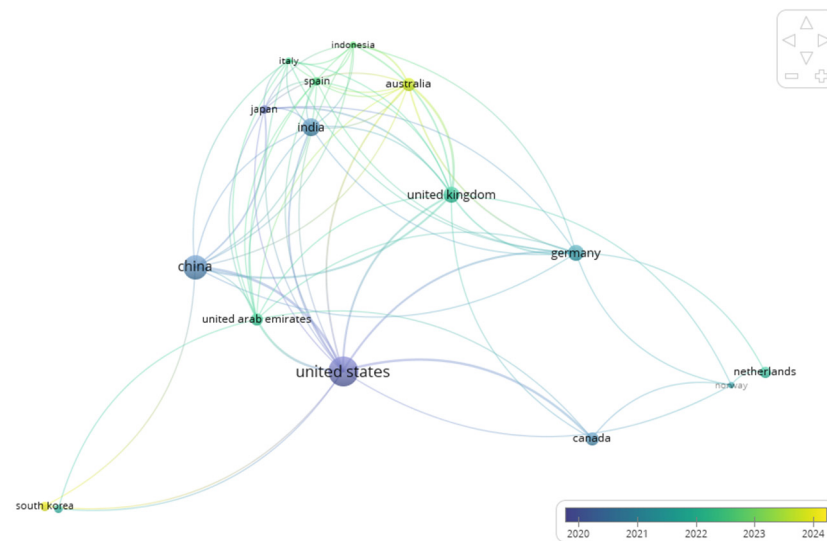
Top sources, authors, and affiliation countries in the corpus (2012–2025). As shown in Table 4A, Lecture Notes in Computer Science (LNCS) anchors the method-building contributions, while TFSC and IJHCI indicate a turn toward managerial and user-experience concerns. Panel A ranks sources by documents (share of  $n = 163$ ); Panel B ranks authors by appearances in the Authors field (share of  $n = 163$ ); Panel C ranks countries by affiliation-string mentions (last token; multi-affiliations counted). LNCS leads among proceedings (5; 3.1%), while TFSC and IJHCI are the most recurrent journals (3 each; 1.8%); the affiliation footprint concentrates in the United States, China, and India. Outlet, authorship, and geography contextualize this pattern. No single venue dominates—proceedings anchor method-building while journals reflect managerial and UX consolidation; a small recurrent core coexists with a long tail of contributors; and the US–China–India triad anchors production with the United Kingdom and Germany as stable secondary nodes. See Table 4A–D.

Proceedings (LNCS) anchor the method-building phase, while TFSC and IJHCI signal consolidation in managerial and HCI journals; no single outlet dominates ( $>5\%$ ).

A small core (e.g., Shin) coexists with a long tail, indicating a distributed authorship base. Table 4B illustrates this author distribution, showing only one author with more than three contributions, and a broad tail of one- or two-time contributors.

Volume rises strongly after 2017 while mid-impact strata thicken after 2019; the US–China–India triad leads, with the UK and Germany as secondary hubs. This is confirmed in Table 4C, which shows a clear dominance of these three countries, followed by the UK and Germany as stable secondary contributors. Taken together, broad outlet dispersion (no venue  $> 5\%$ ), distributed authorship, and a US–China–India footprint indicate an integrative field where explanation design, trust calibration, and exposure-fairness now interlock around e-commerce recommendation problems.

Volumes rise strongly after 2017, while high-citation thresholds are naturally sparse for recent years due to citation latency. Together with Figures 3 and 4, Table 4 provides the production–impact baseline that frames the intellectual structure analysed in the following subsections.



**Figure 4.** Country co-authorship network. Node size reflects total link strength; colors denote SLM clusters; overlay colors indicate average publication year. Source: Scopus (Elsevier); authors' calculations.

#### 4.2. Countries' Co-Authorship

The country co-authorship network (Figure 4) reveals a concentrated collaboration core linking the United States, China, and the United Kingdom, with Germany, India, and Australia acting as high-throughput bridges. Canada and the Netherlands connect this core to additional European and Asia–Pacific ties, while smaller contributors (e.g., Spain, Italy, Japan, United Arab Emirates, South Korea, Norway, Indonesia) attach through one or two dominant partners. Overlay colors indicate a post-2020 intensification of cross-regional collaborations—especially UK/Europe-to-Asia links and transatlantic ties anchored in the US. The modular structure (SLM clusters) separates a North-Atlantic hub from a Europe–Asia–Pacific module, consistent with the field's dual anchoring in information systems/HCI venues and computer science outlets. Overall, output volume and international reach co-evolve: countries with the highest document counts also concentrate Total Link Strength (TLS). At the same time, emerging contributors enter through co-authorship with these hubs rather than via isolated national teams.

Figure 4 shows a modular structure separating a North-Atlantic hub (United States–United Kingdom–Canada–Germany) from a Europe–Asia–Pacific module in which China, India, Australia, and the United Arab Emirates act as high-throughput bridges. The United States ranks first by TLS and partner breadth, indicating the widest dispersion of cross-regional ties. The United Kingdom and Germany sustain dense European linkages while maintaining strong transatlantic connections. China combines a high document volume with broad but slightly more region-weighted ties. Outside the core, Australia and the United Arab Emirates facilitate long-range links to Asia and the Middle East, whereas Japan, Spain, Italy, and Indonesia appear as intensively collaborative satellites connecting through a limited number of dominant partners. The overlay colors suggest a post-2020 intensification of Europe/UK-to-Asia ties and US-anchored transatlantic links, consistent with the field's dual anchoring in information-systems/HCI venues and computer-science outlets.

To complement the structural patterns shown in Figure 4, Table 5 provides a quantitative breakdown of all countries meeting the inclusion thresholds, ranked by Total Link Strength (TLS) and accompanied by their document counts and distinct partner links. TLS captures the cumulative weight of a country’s co-authorship ties and therefore serves as a concise indicator of its collaborative reach within the largest connected set.

**Table 5.** Country co-authorship (thresholds applied), ranked by Total Link Strength (TLS).

Rank	Country	Documents	Links	TLS
1	United States	49	34	51
2	United Kingdom	15	26	36
3	Germany	15	28	32
4	China	34	24	31
5	Australia	10	23	29
6	United Arab Emirates	9	24	27
7	India	18	25	26
8	Japan	4	22	24
9	Indonesia	3	22	24
10	Spain	4	22	22
11	Italy	3	22	22
12	Canada	10	5	8
13	Norway	3	6	6
14	South Korea	5	3	3
15	France	4	3	3

Abbreviations: TLS = Total Link Strength.

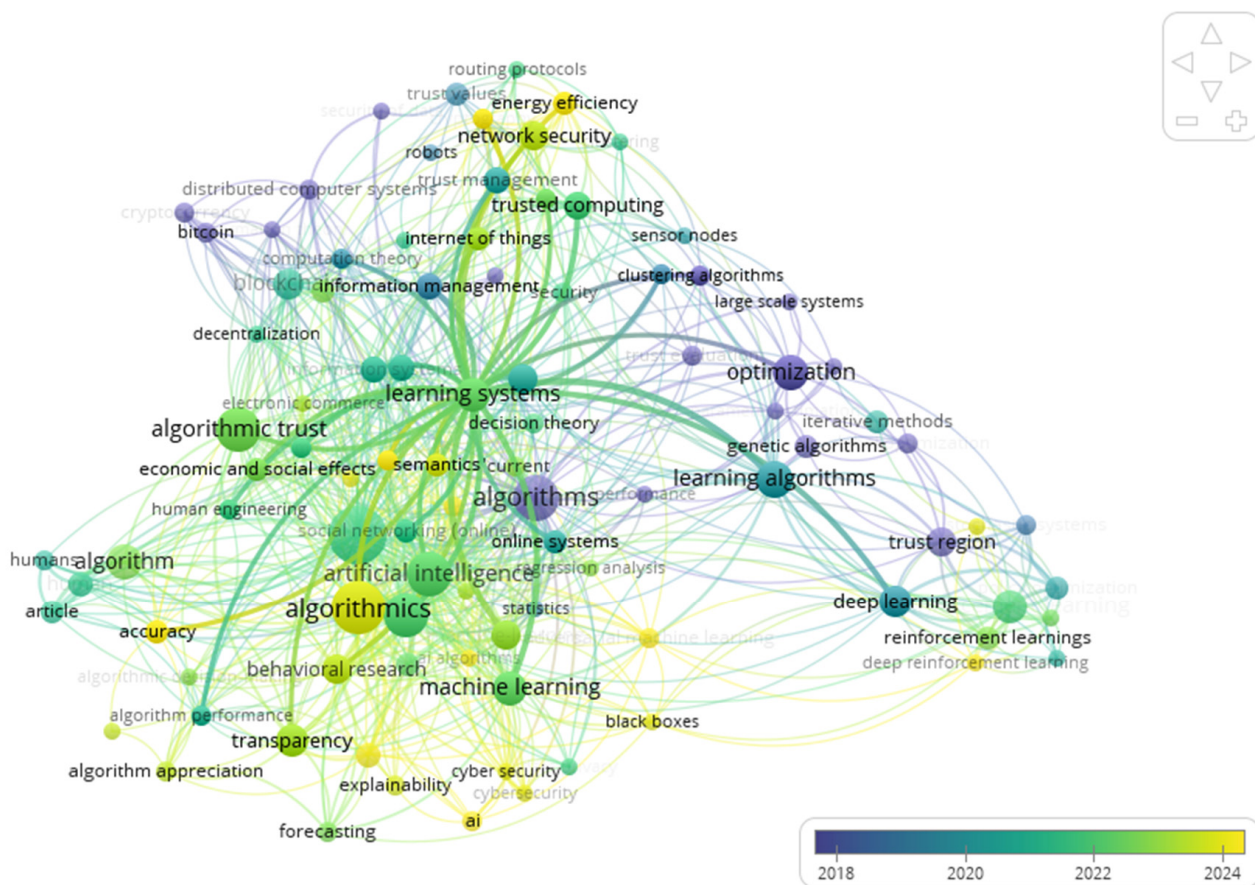
Rankings by TLS align with the visual structure: the United States, United Kingdom, Germany, and China dominate the collaborative core, with Australia, the United Arab Emirates, and India providing high-throughput connections to Asia and the Middle East. Japan, Spain, Italy, and Indonesia—despite smaller outputs—achieve notable partner breadth, consistent with an integration strategy based on a few strong cross-regional ties. Canada, Norway, South Korea, and France meet the participation threshold, but with lower TLS and fewer long-range bridges, indicating more selective international engagement. Together, Figure 4 and Table 5 confirm that collaboration intensity is concentrated in a limited number of hubs, while new or smaller contributors tend to enter via partnerships with those hubs rather than through self-contained national teams.

#### 4.3. Keyword Co-Occurrence (Author Keywords)

Figure 5 visualizes the author-keyword network, revealing a dense conceptual core built around trust, with algorithmic trust, artificial intelligence, transparency, fairness, and explainable AI forming the principal co-occurrence backbone. Machine learning and reinforcement learning connect this backbone to method strands, while recommendation, security, and IoT provide application anchors. A distinct governance–markets satellite links transparency and trust to blockchain, cryptocurrency, and bitcoin, indicating that financial-infrastructure debates have permeated the trust discourse in algorithmic systems. Overall, the map suggests that user-facing trust calibration and system-level accountability (transparency, fairness, explainability) co-evolve and that methods and applications plug into this axis rather than forming isolated islands.

As shown by the overlay colors in Figure 5, early years emphasized explanation and usability, while post-2017 nodes shift toward fairness and accountability vocabulary. In the top-10 by total link strength, four governance-aligned terms—transparency, fairness, explainable AI, and blockchain/cryptocurrency—enter and persist after 2017, i.e., 4/10 items in the post-2017 overlay, evidencing the governance turn. Table 6 ranks the most frequent

keywords by Total Link Strength (TLS), indicating that “trust” and “algorithmic trust” anchor the network’s backbone, while *transparency/fairness/XAI* bridge toward governance themes, and *machine learning/reinforcement learning* connect method strands.



**Figure 5.** Author-keyword co-occurrence network. Node size reflects keyword occurrences; edge thickness indicates co-occurrence strength; colors denote SLM clusters; overlay colors indicate average publication year. Source: Scopus (Elsevier); authors’ calculations. Abbreviations: SLM = smart local moving algorithm.

The presence of blockchain/cryptocurrency/bitcoin among high-TLS keywords signals the migration of trust debates into economic-infrastructure contexts. Rankings by TLS confirm that trust is the hub through which governance (transparency, fairness, explainable AI) and methods (machine learning, reinforcement learning) interconnect with applications (recommendation, IoT, security). Algorithmic trust functions as the pivotal bridge from behavioral constructs to system design. The presence of blockchain, cryptocurrency, and bitcoin among the highest-TLS terms indicates a sustained migration of trust concerns into economic-infrastructure debates. Taken together, Figure 5 and Table 6 suggest that the field’s conceptual gravity centers on trust calibration and accountability by design, with technical methods and application areas attaching to that axis.

Across parameter sweeps, community assignments were stable: the Adjusted Rand Index across runs exceeded 0.80 for the co-citation network and 0.76 for the author-keyword network, suggesting that the reported clusters are unlikely to be artifacts of a single normalization or threshold.

TLS rankings confirm a trust-centered structure in which governance constructs (transparency, fairness, explainable AI) and method families (machine learning, reinforcement learning) connect to application areas (recommendation, IoT, security), with blockchain-related terms signalling the spillover of trust debates into economic-infrastructure domains.

**Table 6.** Author-keyword co-occurrence ranked by Total Link Strength (TLS). Note: Full counting was applied at the paper level. Only keywords with at least three occurrences in the largest connected set were retained.

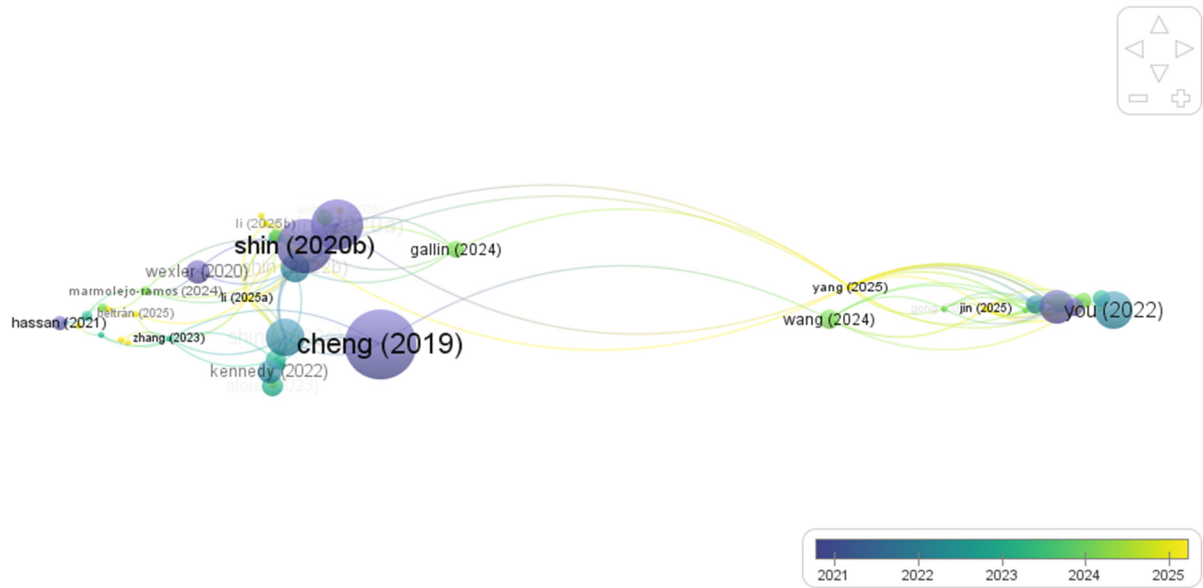
Keyword	Occurrences	Links	TLS
trust	29	15	122
algorithmic trust	16	9	84
artificial intelligence	11	6	54
transparency	8	9	40
algorithms	7	5	37
blockchain	6	4	36
explainable ai	5	6	27
fairness	5	7	27
algorithm	6	4	26
cryptocurrency	4	4	25
machine learning	5	2	22
bitcoin	4	4	21
security	4	5	19
algorithm appreciation	4	4	19
trust management	4	2	17
recommendation	3	2	14
clustering	3	2	14
trust evaluation	3	1	13
IoT	3	4	13
algorithm aversion	3	2	10

Abbreviations: TLS = Total Link Strength; IoT = Internet of Things.

#### 4.4. Bibliographic Coupling of Documents (Frontier Proximity)

Figure 6 visualizes the bibliographic coupling network, which links papers that share references and therefore signals proximity on the research frontier. In this corpus, this proximity concentrates in recent work on recommender-facing trust calibration: explanation design, error tolerance, and performance framing as levers that raise or dampen user reliance. A parallel strand couples journalism/HCI/platform-governance studies through shared references on automation opacity and accountability, tying interface cues to legitimacy claims in commercial settings. Taken together, coupling indicates that the current edge of the field moves from generic “AI trust” toward recommender-specific acceptability: how explanations, error handling, and distributional constraints interact to produce trusted and legitimate ranking outcomes in e-commerce. This extends the co-citation view (intellectual bases) with a frontier view (where new combinations emerge), aligning with our study focus on trust, transparency, fairness, and explainability in retail recommenders.

Table 7 lists the most strongly coupled documents—i.e., those that share the most references with others in the field—highlighting papers that sit at the intersection of design choices (explanations, errors) and trust calibration.



**Figure 6.** Bibliographic coupling network of documents. Node size reflects total coupling strength; edge thickness indicates the number of shared references; colors denote SLM clusters; overlay colors indicate average publication year. Source: Scopus (Elsevier); authors’ calculations. Abbreviations: TCS = Total Coupling Strength; SLM = smart local moving algorithm. Author–year labels denote corpus records, not citations.

**Table 7.** Top coupled documents by Total Coupling Strength (TCS) in the bibliographic-coupling network (threshold:  $\geq 2$  shared references). Counting and parameters: full counting; association-strength normalisation; links formed when two documents share references; link weight = number of shared references; largest connected set retained.

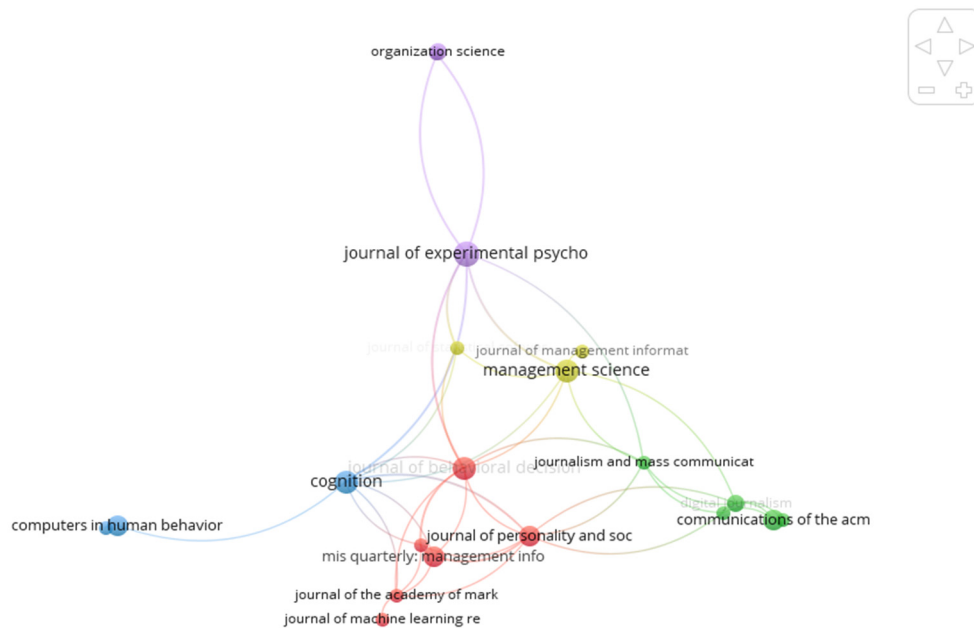
Bibliometric Record Label (First Author–Year–Shortened Title)	Source	Year	Links	TCS
Han (2025)—Trust Formation, Error Impact, and Algorithmic Choices. . .	Information Systems Frontiers	2025	4	16
Efendić (2020)—Slow response times undermine algorithmic advice	Judgment and Decision Making	2020	5	15
Liang (2024)—Trusting an Algorithm Can Be a Smart Move	Journal of Behavioral Decision Making	2024	5	14
Xu (2024)—“Win-win” vs. “sacrifice”: Impact of explanations. . .	Behaviour & Information Technology	2024	5	13
Daschner (2022)—Algorithm aversion? On the influence of errors	Journal of Behavioral Decision Making	2022	4	13
Mahmud (2024)—Decoding algorithm appreciation in everyday AI	Telematics and Informatics Reports	2024	3	10
Jin (2025)—The Influence of ‘Algorithm Aversion’ on Trust. . .	Sustainability	2025	3	9
Yang (2025)—Unveiling users’ algorithm trust dynamics	Information Technology & People	2025	2	6
Wang (2024)—The rationality of explanation on AI trust	Information Processing and Management	2024	1	2
Lee (2024)—Explainable Deep Learning for False Information. . .	Information Systems Research	2024	1	2
Ramrath (2024)—Trust in AI Chatbots: The Perception of Errors	Frontiers in Artificial Intelligence and Applications	2024	1	2

Abbreviations: TCS = Total Coupling Strength; Links = number of distinct partner documents (node degree). Entries in the first column identify bibliometric records from the analysed corpus and are not in-text reference citations.

Rankings and structure jointly show at Table 7 that our niche—trust in e-commerce recommender systems sits at the coupling core: explanation design and error handling are the immediate, designable determinants of reliance; governance and fairness papers remain coupled via shared references, ensuring that the frontier is not only about higher acceptance but also about legitimate distribution of exposure in commercial ranking. These findings delineate the operational space for accountability by design in retail recommender systems, a space that this study directly targets and conceptualizes.

#### 4.5. Co-Citation of Cited References

To identify the intellectual foundations of the field, co-citation analysis was conducted on sources cited together within the corpus. Figure 7 presents the resulting co-citation network, revealing clusters of frequently co-cited references that form the conceptual backbone of algorithmic trust research.



**Figure 7.** Co-citation network of cited sources. Colors denote SLM clusters. Source: Scopus (Elsevier); authors’ calculations. Abbreviations: SLM = smart local moving algorithm.

The top co-cited items (Table 8) cluster around two intellectual pillars: (i) algorithm aversion/appreciation and human–automation reliance (e.g., Dietvorst, Bigman, Burton), and (ii) opacity/governance in algorithmic systems. They explain why trust-calibration levers at the interface (explanation, error handling, controllability) repeatedly co-occur with fairness/accountability concerns at the system level.

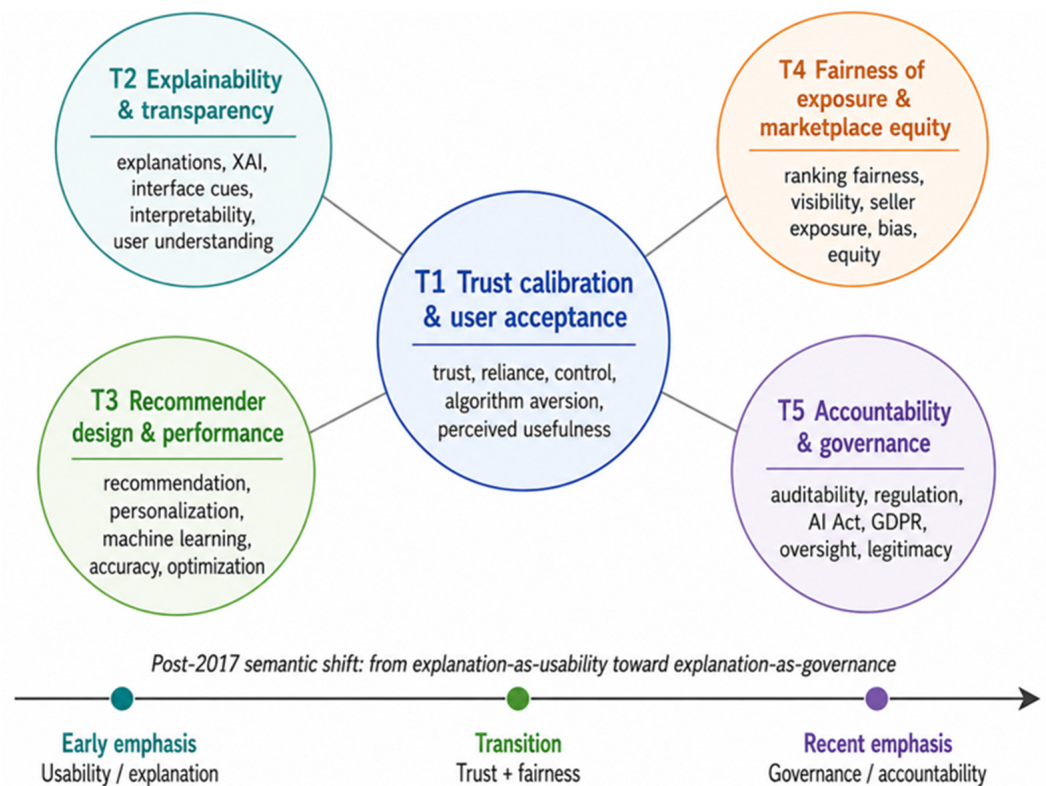
**Table 8.** Top 10 co-cited references by total co-citation link strength (corpus  $n = 163$ ). Counting and parameters: compute co-citation from the Scopus “References” field; link weight = number of within-paper co-citation pairs summed per item (total co-citation link strength); “Times cited in corpus” = number of papers in the corpus that cite the item; inclusion: minimum 2 citations; largest connected set retained.

Rank	Cited Reference	First Author	Year	Times Cited in Corpus ( $n$ of 163)	Total Co-Citation Link Strength
1	[43]	Castelo	2019	10	86
2	[38]	Ananny	2018	9	79
3	[28]	Alexander	2018	6	54
4	[76]	Dietvorst	2015	6	52
5	[14]	Burrell	2016	6	52
6	[30]	Bigman	2018	5	45
7	[73]	Beer	2017	5	45
8	[70]	Burton	2020	5	43
9	[77]	Cadario	2021	3	27
10	[31]	Dietvorst	2018	3	27

#### 4.6. BERTopic-Derived Thematic Structure

To complement the bibliometric networks, BERTopic was used as a semantic layer for identifying the dominant thematic structures in the corpus. The resulting topics should not be interpreted as statistically independent causal dimensions, but as recurring semantic

concentrations that help explain how the field connects trust, explainability, fairness, user interaction, and governance. Figure 8 summarises the five retained topic groups and visualises their temporal interpretation as a post-2017 shift from explanation- and usability-oriented concerns toward fairness, accountability, and governance.



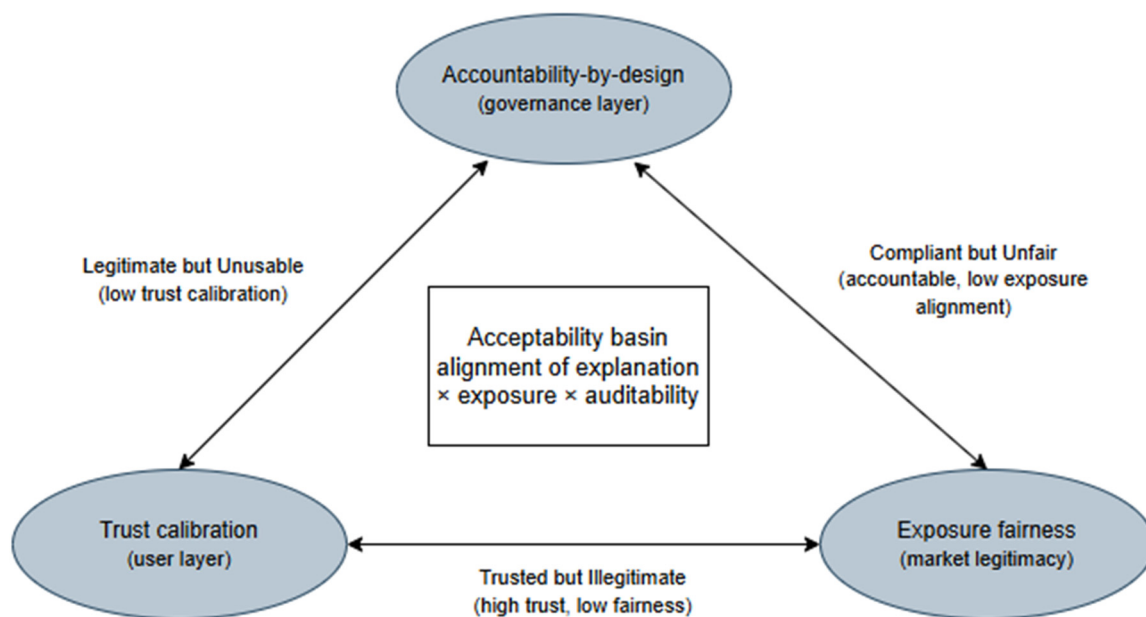
**Figure 8.** BERTopic-derived thematic structure and post-2017 semantic shift in the corpus. Source: Authors’ illustration based on BERTopic modelling of the Scopus corpus ( $n = 163$ ; 2012–2025).

Figure 8 visualises the BERTopic-derived thematic structure of the corpus. The topics indicate a semantic movement from explanation- and usability-oriented concerns toward fairness, exposure, governance, and accountability. Rather than serving as a confirmatory model, the figure provides an interpretive visual aid that supports the synthesis developed in Section 4.7 and the Acceptance Triangle introduced in Section 5.

#### 4.7. Synthesis Across Layers

Read together, the keywords, bibliographic-coupling, co-citation, and topic-modelling lenses support an interpretive synthesis rather than a causal claim. The keyword map (Figure 5; Table 6) centres the field on trust and algorithmic trust, linking explanation to fairness and accountability, while method strands attach via machine-learning terms and economic-infrastructure keywords. The bibliographic-coupling network (Figure 6; Table 7) places the recent frontier at the interface, where explanation type, error transparency, and controllability operate as immediate levers of user reliance in recommender use. The co-citation base (Figure 7; Table 8) anchors this focus in two pillars: human–automation reliance, including algorithm aversion and appreciation, and accounts of opacity and governance in algorithmic systems. From this convergence, we propose a cross-layer interpretive principle: acceptable outcomes in retail recommender systems are more plausibly understood when user-level trust calibration, marketplace-level exposure equity, and governance-level accountability by design are considered jointly rather than in isolation.

This reframes the target from trust alone to algorithmic acceptability and sets the conceptual foundation for the Acceptance Triangle and the managerial playbook developed in the Discussion (Figure 9; Section 5).



**Figure 9.** The Acceptance Triangle: alignment of trust calibration (user layer), exposure equity (market layer), and accountability-by-design (governance layer). Source: Authors' illustration based on the synthesis of keyword co-occurrence, bibliographic coupling, co-citation, and topic-modelling results (Scopus corpus,  $n = 163$ ; 2012–2025).

## 5. Discussion

The Acceptance Triangle is developed here as a conceptual synthesis grounded in the patterns observed across the bibliometric, coupling, co-citation, and topic-modelling analyses. It is not presented as a causal model directly proven by bibliometric evidence. Rather, it organises three recurrent layers identified in the mapped literature: user-facing trust calibration, marketplace-level exposure equity, and governance-level accountability. This distinction is important because bibliometric evidence can reveal conceptual proximity, thematic consolidation, and research trajectories, but the translation of these patterns into a framework remains an interpretive theoretical contribution.

The findings extend traditional IS acceptance models by suggesting that perceived usefulness and trust outcomes in e-commerce recommender systems should be interpreted alongside fairness and accountability considerations. The proposed cross-layer framework includes: (i) interface-level trust calibration, including explanation, error transparency, and controllability; (ii) exposure equity at the marketplace level; and (iii) governance-layer accountability by design. In EU markets, such alignment is consistent with regulatory expectations concerning transparency and auditability [28,71]. Rather than maximising trust alone, the relevant design goal becomes acceptability, understood as the alignment of user-facing calibration, distributional fairness, and traceable system accountability. At the recommender-system level, transparency mechanisms can strengthen consumer trust when they are connected to perceived usefulness, explanation quality, and user-facing control rather than treated as isolated technical disclosures [32].

Figure 9 helps reconcile the mixed findings by making three recurrent failure modes visible that our maps anticipate. Relatedly, recent strategic communication research frames interface and policy configurations as points of potential algorithmic manipulation, with measurable cognitive effects for both users and markets [93,94]. Content governance

debates also demonstrate how speech norms and enforcement architectures intersect with user trust in platform curation [95,96].

The first pattern, *Trusted but Illegitimate*, emerges when interface design increases reliance, but ranking policies allocate visibility inequitably. While users may comply, marketplace actors may perceive systemic unfairness [97]. Second, *Compliant but Unfair*: formal accountability artifacts exist, yet exposure remains misaligned with stated objectives, an issue consistent with governance-style remedies that do not bind distributional outcomes [98,99]. Third, *Legitimate but Unusable*: distributional constraints and accountability are present, but explanation and controllability fail to calibrate expectations, depressing sustained use or creating pockets of over-reliance [100].

These patterns mirror the bibliographic-coupling frontier on explanation design, error handling, and user control [101–103], and the topic overlays that link explanation to fairness-of-exposure and accountability after 2017, as reported in Section 3 [104,105]. Segment-level expectation heterogeneity underscores the need for explicit preference elicitation [106]. Technology readiness systematically conditions acceptance. The TRI 2.0 index captures predispositions that moderate trust and reliance [107]. Two testable statements follow. S1 (Cross-layer necessity): Trust gains not coupled with exposure-fairness constraints and auditability raise short-term reliance but do not achieve acceptability in retail recommenders [108]. S2 (Alignment over magnitude): Marginal improvements on any single layer yield diminishing returns unless alignment with the other two layers is preserved or improved [109,110].

Together, S1–S2 summarize why single-layer optimizations in prior work (e.g., transparency only, or governance only) produce heterogeneous effects across contexts [111,112]. Both statements are field-testable with concurrent A/B manipulations of explanation, error handling, and controllability under explicit exposure constraints, evaluated against a composite acceptability score rather than accuracy alone [113].

The co-citation structure anchors three intellectual pillars: appropriate reliance in automation and trust, transparency/explanation in recommenders, and fairness/accountability and principles [114,115]. The bibliographic-coupling front concentrates on explanation type, error transparency, and controllability as near-term design levers that influence reliance, consistent with algorithm-aversion/appreciation dynamics [116,117].

Finally, the topic-modelling layer, interpreted together with the lexicon-aligned keyword overlay, supports the reading of a post-2017 governance shift: explanation discourse increasingly co-occurs with fairness-of-exposure and accountability terminology, shifting the centre of gravity from explanation-as-usability toward explanation-as-governance [118,119]. This trajectory also aligns with privacy and risk salience trends that make risk communication integral to acceptance: large-scale evidence shows sustained public concern over data practices [120], while governance work emphasizes traceability and obligation [121].

Crisis periods amplified this salience, expanding surveillance rationales and recalibrating expectations of data use. These trajectories align with accounts of a routine “surveillance society” in which monitoring is infrastructural rather than exceptional [122,123]. In our reading of the frontier, explanation without error transparency and control tends to miscalibrate reliance; governance without exposure constraints risks performative compliance; and fairness-only adjustments can remain invisible to users unless articulated and auditable [124].

Operationally, it translates the Acceptance Triangle into nine design levers with testable KPIs (Table 9): explanation type, error transparency, controllability, objective disclosure, exposure constraints, preference elicitation, risk communication, auditability, and multi-objective evaluation across fairness, transparency, and accountability dimensions [125,126].

This framing accommodates interface modalities where presentational choices can amplify or backfire. For example, immersive AR elements may lift perceived quality yet hurt calibration if explanation and control are not co-designed [127].

**Table 9.** Trust UX playbook: levers, managerial hypotheses, and suggested KPIs.

Proposition ID	Lever	Managerial Hypothesis	Suggested KPI(s)
P1	Type of explanation	Counterfactual/rationale-style explanations outperform feature lists for calibrated reliance in higher-stakes or higher-cost purchases.	Calibrated reliance; post-explanation conversion; reliance error rate
P2	Error transparency	Rapid error acknowledgment and corrective action increase post-error trust retention relative to silent correction.	Post-error trust retention; complaint rate; time-to-recovery
P3	Controllability (user tweaks)	Simple preference controls reduce algorithm aversion and increase sustained usage without harming accuracy.	24 h/30 d retention; control usage rate; accuracy delta
P4	Objective disclosure	Explicit system objectives combined with fairness constraints raise perceived legitimacy among merchants and users.	Perceived legitimacy; merchant NPS; fairness-metric compliance
P5	Exposure constraints	Amortized exposure equity reduces seller complaints with negligible CTR loss vs. unconstrained ranking.	Seller complaint rate; exposure equity, CTR delta
P6	Preference elicitation (cold-start)	Richer initial elicitation reduces early mismatch churn and lifts first-week conversion.	Early churn, first-week conversion, mismatch tickets
P7	Risk communication	Clear articulation of privacy/cost trade-offs increases trust, without over-reliance, vs. minimal notices.	Trust score; over-reliance index; opt-out rate
P8	Auditability (logs & traceability)	Providing audit trails and explainability metrics raises B2B platform trust and reduces dispute time.	B2B trust score; dispute-resolution time; audit completeness
P9	Multi-objective evaluation	A composite KPI (accuracy + fairness + complaints) predicts merchant retention better than accuracy-only baselines.	Composite score; merchant retention; NDCG vs. composite R <sup>2</sup>

Abbreviations: KPI = key performance indicator; NPS = Net Promoter Score; CTR = click-through rate; NDCG = normalized discounted cumulative gain; R<sup>2</sup> = coefficient of determination; calibrated reliance = alignment between user reliance and actual system performance; exposure equity = distributional fairness of visibility over time.

Re-identification results on large, sparse datasets underscore why mere “anonymisation” is a weak guarantee absent governance and auditability [128]. It also accommodates infrastructure-level enablers of accountability such as auditable logs and tamper-evident records, including blockchain-style mechanisms in e-commerce contexts, provided they are tied to exposure policies and objective disclosure rather than used as standalone trust signals [129].

Two caveats temper interpretation. First, our evidence is single-database (Scopus) and inherits coverage and recency biases for 2023–2025 cohorts. Second, semantic overlays depend on thesaurus choices and thresholds; we provide full materials for reuse. Beyond our scope, sector-specific governance and security literature outline contextual pressures that can shape acceptance paths—e.g., privacy governance and contextual integrity, cybersecurity obligations in critical sectors, and platform-level manipulation or UX dark patterns. However, these should be integrated only when they instrumentally alter explanation, exposure, or auditability in the given retail setting [130].

Broader governance and cybersecurity literature are therefore relevant only insofar as they directly affect explanation, exposure, auditability, or dispute resolution in e-commerce recommender systems. For the present study, these literatures are treated as a contextual boundary condition rather than as an independent analytical domain.

### 5.1. Managerial Implications: The Trust UX Playbook

The convergence of maps and overlays supports steering systems into the basin of acceptability—where explanation design, exposure constraints, and auditability are aligned—thereby grounding the Trust UX Playbook (Table 9) as a practical framework for aligning technical, user, and governance objectives.

The convergence of our science maps and topic overlays helps steer system design into what we define as the basin of acceptability, a space where explanation, exposure, and auditability are not treated as separate concerns but jointly optimized. This perspective grounds the Trust UX Playbook (Table 9) as a practical framework for aligning user experience with broader governance and market legitimacy goals. Each of the nine levers, from

explanation type to multi-objective evaluation, addresses a specific tension observed in prior research. For instance, error transparency and controllability are critical for sustaining user trust beyond initial engagement, while exposure constraints and auditability speak to the fairness and traceability of algorithmic outputs. The hypotheses and KPIs listed offer a structured agenda for experimentation, enabling iterative improvements across interface, policy, and infrastructure levels.

To support holistic evaluation, we introduce the Composite Acceptability Score (CAS) as an exploratory managerial illustration that integrates accuracy, fairness of exposure, and complaint reduction. Rather than replacing accuracy, CAS complements it by showing how technical performance may be evaluated alongside distributional and accountability-related indicators. The proposed weighting structure is not presented as empirically validated or universally applicable. Instead, it offers a transparent starting configuration that future studies should recalibrate using real-world platform data, stakeholder priorities, sensitivity analysis, and sector-specific risk profiles.

Together, P1–P9 offer a design and evaluation playbook that can be adapted to platform-specific contexts, policy regimes, and user segments. As retail platforms face growing scrutiny over transparency and fairness, these levers provide not only technical direction but also normative grounding for accountable algorithmic design.

## 5.2. Limitations and Future Research

This study relies on one index (Scopus) and English-language records; recency effects apply to 2023–2025 cohorts. Semantic overlays depend on thesaurus choices and frequency thresholds. We mitigate these risks through transparent parameters, robustness checks (full vs. fractional counting; threshold sweeps), and stability metrics (ARI). All queries, exports, thesaurus files, and analysis scripts are available in an anonymized repository for full replication. Sector-specific governance and security studies may shape acceptance paths in domain-contingent ways, for example, privacy governance and contextual integrity, cybersecurity obligations in critical sectors, platform manipulation/dark patterns, and design-oriented agendas on manipulation in generative AI. These streams should be integrated only when they instrumentally alter explanation, exposure, or auditability in the given retail setting. Future work can extend the corpus to multiple indices, triangulate overlays with manual topic audits, and test the Acceptance Triangle via preregistered field experiments that report accuracy alongside distributional and governance metrics.

Future research should also validate the CAS weighting structure empirically. The illustrative weights assigned to accuracy, exposure equity, and complaint reduction should be tested through platform-level A/B experiments, multi-criteria decision analysis, expert elicitation, stakeholder calibration, and longitudinal validation against outcomes such as user retention, merchant complaints, perceived legitimacy, and dispute-resolution time. Such studies could determine whether the proposed weights remain stable across platform types or need to be adapted to different sectors, jurisdictions, and risk contexts.

External validity may vary across jurisdictions, platform types, and regulatory regimes. Future studies should therefore examine whether the Acceptance Triangle operates differently in marketplaces with stronger transparency obligations, stricter consumer-protection rules, or higher exposure-related risks for sellers.

## 6. Conclusions

Over the last decade, scholarship on trust in algorithmic decision-making within e-commerce recommender systems has shifted from a usability-centred agenda toward a governance-aware and market-sensitive understanding of acceptability. Across the bibliometric maps, trust calibration, exposure equity, and accountability by design appear

as increasingly connected but still analytically distinct themes. On this basis, the study proposes the Acceptance Triangle as a conceptual framework for interpreting algorithmic acceptability in retail recommender systems. The Composite Acceptability Score (CAS) is introduced as an exploratory managerial illustration rather than as a validated decision rule. For transparency, the illustrative specification is  $CAS = 0.40 \times \text{Accuracy} + 0.35 \times \text{Exposure equity} + 0.25 \times (1 - \text{Normalized complaints})$ , with all components min–max-scaled to  $[0, 1]$ . Future empirical studies should validate, recalibrate, or replace these weights using real-world platform data.

Science-mapping of the Scopus corpus (2012–2025;  $n = 163$ ) shows steady output growth with a post-2017 inflection (Figure 2) and a thickening mid-impact stratum after 2019 (Figure 3), while collaboration consolidates around a transatlantic–Asia-Pacific core (Figure 4). Conceptually, author-keyword structures place trust and algorithmic trust at the center, linking explanation, fairness, and auditability with application strands in recommendation and commerce, and with infrastructural debates signaled by blockchain terms (Figure 5). On the frontier, bibliographic coupling concentrates on explanation type, error transparency, and controllability as immediate levers of reliance in recommender use (Figure 6).

These lenses motivate a reframing: the relevant target is not trust per se but acceptability, which can be more comprehensively interpreted when user-level trust calibration, marketplace-level exposure equity, and governance-level accountability are considered together. The Acceptance Triangle (Figure 9) conceptualises this alignment and helps explain heterogeneous findings in prior work: trusted yet illegitimate allocations, compliant yet unfair outcomes, or legitimate configurations that remain unusable. Operationally, UI-only or accuracy-only optimization should be replaced by multi-objective steering evaluated with CAS.

Managerially, the path is actionable. Table 9 translates the Triangle into nine levers—explanation type, error transparency, controllability, objective disclosure, exposure constraints, preference elicitation, risk communication, auditability, and multi-objective evaluation—each accompanied by testable propositions and KPIs. Do not adopt fairness constraints that are invisible to users or unverifiable, and do not claim accountability without logs, metrics, and dispute traceability. In retail settings where ranking allocates scarce visibility, aligning these layers is a requirement for durable legitimacy and performance. This model and metric provide a reproducible framework for aligning UX design, fairness constraints, and governance in algorithmic recommendation systems.

A single database (Scopus) and English-language records, recency effects for 2023–2025 cohorts, and dependence of semantic overlays on thesaurus choices constrain generalization. Mitigation stems from transparent parameters and reproducibility materials. Nonetheless, the directional pattern is clear: since 2017, explanation discourse has increasingly co-occurred with exposure-fairness and accountability terminology, and the research frontier has shifted from “making systems usable” to “making systems acceptable”.

The next step is empirical. Future studies should run field experiments that co-manipulate explanation, error handling, and user control under explicit exposure policies and auditable objectives; compare CAS against accuracy-only KPIs for predicting merchant retention and user welfare; and test sectoral boundary conditions where governance obligations alter acceptance paths.

**Author Contributions:** Conceptualization, M.G. and M.P.B.; methodology, M.G. and M.P.B.; software, M.G.; validation, M.G. and M.P.B.; formal analysis, M.G.; investigation, M.G.; resources, M.G., A.T., and M.P.B.; data curation, M.G.; writing—original draft preparation, M.G.; writing—review and editing, M.G., A.T. and M.P.B.; visualization, M.G.; supervision, M.P.B.; project administration, M.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data and analysis materials supporting the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflicts of interest. Amir Topalović is affiliated with Aisma Srl; however, the company provided no funding for this study and had no role in the research or the decision to publish the results.

## References

- Jarrahi, M.H. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Bus. Horiz.* **2018**, *61*, 577–586. [CrossRef]
- Acemoglu, D.; Restrepo, P. Automation and new tasks: How technology displaces and reinstates labor. *J. Econ. Perspect.* **2019**, *33*, 3–30. [CrossRef]
- Kellogg, K.C.; Valentine, M.A.; Christin, A. Algorithms at work: The new contested terrain of control. *Acad. Manag. Ann.* **2020**, *14*, 366–410. [CrossRef]
- Bankins, S.; Weaver, A.; Marrone, M.; Restubog, S.L.D.; Woo, S.E. Artificial intelligence at work: A phenomenon-based interdisciplinary review. *Inf. Manag.* **2026**, *63*, 104272. [CrossRef]
- Deranty, J.P.; Corbin, T. Artificial intelligence and work: A critical review of recent research from the social sciences. *AI Soc.* **2024**, *39*, 675–691. [CrossRef]
- Raisch, S.; Krakowski, S. Artificial intelligence and management: The automation–augmentation paradox. *Acad. Manag. Rev.* **2021**, *46*, 192–210. [CrossRef]
- Rahwan, I.; Cebrian, M.; Obradovich, N.; Bongard, J.; Bonnefon, J.-F.; Breazeal, C.; Crandall, J.W.; Christakis, N.A.; Couzin, I.D.; Jackson, M.O.; et al. Machine behaviour. *Nature* **2019**, *568*, 477–486. [CrossRef]
- Feuerriegel, S.; Hartmann, T.; Janiesch, C.; Zschech, P. Generative AI. *Bus. Inf. Syst. Eng.* **2024**, *66*, 111–126. [CrossRef]
- Mittelstadt, B.; Allo, P.; Taddeo, M.; Wachter, S.; Floridi, L. The ethics of algorithms: Mapping the debate. *Big Data Soc.* **2016**, *3*, 2053951716679679. [CrossRef]
- Binns, S.; Veale, M. Is that your final decision? Multi-stage profiling, selective effects, and the GDPR. *Int. Data Priv. Law* **2021**, *11*, 253–270. [CrossRef]
- Benjamin, R. *Viral Justice: How We Grow the World We Want*; Princeton University Press: Princeton, NJ, USA, 2022.
- Eubanks, V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*; St. Martin's Press: New York, NY, USA, 2018.
- Ananny, M. Toward an ethics of algorithms: Convening, observation, probability, and timeliness. *Sci. Technol. Hum. Values* **2016**, *41*, 93–117. [CrossRef]
- Burrell, J. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data Soc.* **2016**, *3*, 2053951715622512. [CrossRef]
- Seaver, N. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data Soc.* **2017**, *4*, 2053951717738104. [CrossRef]
- Gillespie, T. The relevance of algorithms. In *Media Technologies: Essays on Communication, Materiality, and Society*; Gillespie, T., Boczkowski, P., Foot, K., Eds.; MIT Press: Cambridge, MA, USA, 2014; pp. 167–194.
- Amoore, M. Cloud geographies: Computing, data, sovereignty. *Prog. Hum. Geogr.* **2018**, *42*, 4–24. [CrossRef]
- Trielli, D.; Diakopoulos, N. Algorithmic agenda-setting: The shape of search media during the 2020 US election. *Int. Symp. Online Journal.* **2022**, *12*, 45–70.
- Zhang, S.; Chen, X. Explainable recommendation: A survey and new perspectives. *Found. Trends Inf. Retr.* **2020**, *14*, 1–101. [CrossRef]
- Morley, J.; Elhalal, A.; Garcia, F.; Kinsey, L.; Mökander, J.; Floridi, L. Ethics as a service: A pragmatic operationalisation of AI ethics. *Minds Mach.* **2021**, *31*, 239–256. [CrossRef]
- Araujo, T.; Helberger, N.; Kruike-meier, S.; de Vreese, C.H. In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI Soc.* **2020**, *35*, 611–623. [CrossRef]
- Dwivedi, Y.K.; Jeyaraj, A.; Hughes, L.; Davies, G.H.; Ahuja, M.; Albashrawi, M.A.; Al-Busaidi, A.S.; Al-Sharhan, S.; Al-Sulaiti, K.I.; Altinay, L.; et al. Real impact: Challenges and opportunities in bridging the gap between research and practice—Making a difference in industry, policy, and society. *Int. J. Inf. Manag.* **2024**, *78*, 102750. [CrossRef]

23. Martínez-López, F.J.; Rodríguez-Ardura, I.; Gázquez-Abad, J.C.; Sánchez-Franco, M.J.; Cabal, C.C. Psychological elements explaining the consumer's adoption and use of a website recommendation system: A theoretical framework proposal. *Internet Res.* **2010**, *20*, 316–341. [[CrossRef](#)]
24. Voria, G.; Hadjiloucas, S.; Serrano, M.; Tselentis, G. Fairness-aware practices from developers' perspective: A survey. *Inf. Softw. Technol.* **2025**, *182*, 107710. [[CrossRef](#)]
25. Obermeyer, Z.; Powers, B.; Vogeli, C.; Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **2019**, *366*, 447–453. [[CrossRef](#)]
26. Townsend, B.; Parnell, K.J.; Yaman, S.G.; Nemirovsky, G.; Calinescu, R. Normative conflict resolution through human–autonomous agent interaction. *J. Responsible Technol.* **2025**, *21*, 100114. [[CrossRef](#)]
27. Cajueiro, D.O.; Celestino, V.R.R. A comprehensive review of Artificial Intelligence regulation: Weighing ethical principles and innovation. *J. Econ. Technol.* **2026**, *4*, 77–91. [[CrossRef](#)]
28. Alexander, V. Why trust an algorithm? Performance, risk, and trust. *Comput. Hum. Behav.* **2018**, *89*, 279–288. [[CrossRef](#)]
29. Glikson, E.; Woolley, A.W. Human trust in artificial intelligence: Review of empirical research. *Acad. Manag. Ann.* **2020**, *14*, 627–660. [[CrossRef](#)]
30. Bigman, Y.E.; Gray, K. People are averse to machines making moral decisions: The case of autonomous vehicles. *Cognition* **2018**, *181*, 21–34. [[CrossRef](#)]
31. Dietvorst, B.J.; Simmons, J.P.; Massey, C. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Manag. Sci.* **2018**, *64*, 1155–1170. [[CrossRef](#)]
32. Li, Y.; Deng, X.; Hu, X.; Liu, J. The effects of e-commerce recommendation system transparency on consumer trust: Exploring parallel multiple mediators and a moderator. *J. Theor. Appl. Electron. Commer. Res.* **2024**, *19*, 2630–2649. [[CrossRef](#)]
33. Robinson, S.C. Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technol. Soc.* **2020**, *63*, 101421. [[CrossRef](#)]
34. Eidelson, B. Patterned inequality, compounding injustice, and algorithmic prediction. *Am. J. Law Equal.* **2021**, *1*, 252–276. [[CrossRef](#)]
35. Marjanovic, O.; Cecez-Kecmanovic, D.; Vidgen, R. Theorising algorithmic justice. *Eur. J. Inf. Syst.* **2021**, *31*, 269–287. [[CrossRef](#)]
36. Moreno-Sánchez, P.A.; Del Ser, J.; van Gils, M.; Hernesniemi, J. A design framework for operationalizing trustworthy artificial intelligence in healthcare: Requirements, tradeoffs and challenges for its clinical adoption. *Inf. Fusion* **2026**, *127*, 103812. [[CrossRef](#)]
37. Kyung, N.; Kwon, H.E. Rationally trust, but emotionally? The roles of cognitive and affective trust in laypeople's acceptance of AI for preventive care operations. *Prod. Oper. Manag.* **2022**, 1–20. [[CrossRef](#)]
38. Ananny, M. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media Soc.* **2018**, *20*, 973–989. [[CrossRef](#)]
39. Birhane, A. Algorithmic injustice: A relational ethics approach. *Patterns* **2021**, *2*, 100205. [[CrossRef](#)]
40. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A survey of methods for explaining black box models. *ACM Comput. Surv.* **2018**, *51*, 93. [[CrossRef](#)]
41. Rudin, C. Stop explaining black-box machine learning models for high-stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. [[CrossRef](#)]
42. Biega, A.J.; Gummadi, K.P.; Weikum, G. Equity of attention: Amortizing individual fairness in rankings. In Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '18), Ann Arbor, MI, USA, 8–12 July 2018; pp. 405–414. [[CrossRef](#)]
43. Castelo, N.; Bos, M.W.; Lehmann, D.R. Task-dependent algorithm aversion. *J. Mark. Res.* **2019**, *56*, 809–825. [[CrossRef](#)]
44. Saxena, N.A.; Huang, K.; DeFilippis, E.; Radanovic, G.; Parkes, D.C.; Liu, Y. How do fairness definitions fare? Examining public attitudes towards algorithmic definitions of fairness. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, Honolulu, HI, USA, 27–28 January 2019; pp. 99–106. [[CrossRef](#)]
45. Gray, C.M.; Kou, Y.; Battles, B.; Hoggatt, J.; Toombs, A.L. The dark (patterns) side of UX design. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–14. [[CrossRef](#)]
46. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). In *Official Journal of the European Union*; Publications Office of the European Union: Luxembourg, 2024. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689> (accessed on 21 May 2026).
47. Achuthan, K.; Ramanathan, S.; Raman, R. Securing the metaverse: Machine learning-based perspectives on risk, trust, and governance. *Int. J. Inf. Manag. Data Insights* **2025**, *5*, 100356. [[CrossRef](#)]
48. Klenk, M. Ethics of generative AI and manipulation: A design-oriented research agenda. *Ethics Inf. Technol.* **2024**, *26*, 9. [[CrossRef](#)]
49. Mann, M.; Matzner, T. Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. *Big Data Soc.* **2019**, *6*, 2053951719895805. [[CrossRef](#)]

50. Khademi, A.; Wang, L.; Tschantz, A.; Weld, D.S. Fairness in algorithmic decision making: An excursion through the lens of causality. In Proceedings of the World Wide Web Conference, San Francisco CA USA, 13–17 May 2019; pp. 2907–2914. [CrossRef]
51. Mittelstadt, B.D. Principles alone cannot guarantee ethical AI. *Nat. Mach. Intell.* **2019**, *1*, 501–507. [CrossRef]
52. Ghasemaghaei, M.; Kordzadeh, N. Understanding how algorithmic injustice leads to making discriminatory decisions: An obedience to authority perspective. *Inf. Manag.* **2024**, *61*, 103921. [CrossRef]
53. Goodman, B.; Flaxman, S. European Union regulations on algorithmic decision-making and a ‘right to explanation’. *AI Mag.* **2017**, *38*, 50–57. [CrossRef]
54. Berente, N.; Seidel, J.; Safadi, A.; Lyytinen, A. Managing artificial intelligence. *MIS Q.* **2021**, *45*, 1433–1450. [CrossRef]
55. European Union Agency for Cybersecurity (ENISA). NIS 360°: Cybersecurity in Critical Sectors. 2024. Available online: <https://www.enisa.europa.eu> (accessed on 25 March 2026).
56. Kamara, I.; Kosta, E. Do Not Track initiatives: Regaining the lost user control. *Int. Data Priv. Law* **2016**, *6*, 276–290. [CrossRef]
57. Leese, M. The new profiling: Algorithms, black boxes, and the failure of anti-discriminatory safeguards in the European Union. *Secur. Dialogue* **2014**, *45*, 494–511. [CrossRef]
58. Schreurs, W.; Hildebrandt, M.; Kindt, E.; Vanfleteren, M. The role of data protection law and non-discrimination law in group profiling in the private sector. In *Profiling the European Citizen*; Springer: Dordrecht, The Netherlands, 2008; pp. 241–270. [CrossRef]
59. Lyon, D. Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data Soc.* **2014**, *1*, 2053951714541861. [CrossRef]
60. Bankins, S.; Formosa, P. The ethical implications of artificial intelligence for meaningful work. *J. Bus. Ethics* **2023**, *185*, 725–740. [CrossRef]
61. Perdana, A.; Arifin, S.; Quadrianto, N. Algorithmic trust and regulation: Governance, ethics, legal, and social implications blueprint for Indonesia’s central banking. *Technol. Soc.* **2025**, *81*, 102838. [CrossRef]
62. Jungherr, A.; Rauchfleisch, A. Artificial Intelligence in deliberation: The AI penalty and the emergence of a new deliberative divide. *Gov. Inf. Q.* **2025**, *42*, 102079. [CrossRef]
63. Chen, R.; Zolbanin, H.M.; Gentina, E. Pseudo-community trust and member self-disclosure. *Inf. Manag.* **2024**, *61*, 104001. [CrossRef]
64. Hou, T.-Y.; Tseng, Y.-C.; Yuan, C.W. Is this AI sexist? The effects of a biased AI’s anthropomorphic appearance and explainability on users’ bias perceptions and trust. *Int. J. Inf. Manag.* **2024**, *74*, 102775. [CrossRef]
65. Toumaj, S.; Heidari, A.; Navimipour, N.J. Leveraging explainable artificial intelligence for transparent and trustworthy cancer detection systems. *Artif. Intell. Med.* **2025**, *169*, 103243. [CrossRef]
66. Acquisti, A.; Brandimarte, L.; Loewenstein, G. Privacy and human behaviour in the age of information. *Science* **2015**, *347*, 509–514. [CrossRef]
67. Hoff, K.A.; Bashir, M. Trust in automation: Integrating empirical evidence to inform policy and design. *Hum. Factors* **2015**, *57*, 407–434. [CrossRef]
68. Tintarev, N.; Masthoff, J. Evaluating the effectiveness of explanations for recommender systems. *User Model. User-Adapt. Interact.* **2012**, *22*, 399–439. [CrossRef]
69. Gedikli, F.; Jannach, D.; Ge, M. How should I explain? A comparison of different explanation types for recommender systems. *Int. J. Hum.-Comput. Stud.* **2014**, *72*, 367–382. [CrossRef]
70. Burton, J.W.; Stein, M.-K.; Jensen, T.B. A systematic review of algorithm aversion in augmented decision making. *J. Behav. Decis. Mak.* **2020**, *33*, 220–239. [CrossRef]
71. Acharya, N.; Sassenberg, A.-M.; Soar, J. Consumers’ behavioural intentions to reuse recommender systems: Assessing the effects of trust propensity, trusting beliefs, and perceived usefulness. *J. Theor. Appl. Electron. Commer. Res.* **2023**, *18*, 55–78. [CrossRef]
72. Parasuraman, A.; Colby, C.L. An updated and streamlined technology readiness index: TRI 2.0. *J. Serv. Res.* **2015**, *18*, 59–74. [CrossRef]
73. Beer, D. The social power of algorithms. *Inf. Commun. Soc.* **2017**, *20*, 1–13. [CrossRef]
74. Wang, S.; Zhang, X.; Wang, Y.; Ricci, F. Trustworthy recommender systems. *ACM Trans. Intell. Syst. Technol.* **2024**, *15*, 84. [CrossRef]
75. Gombar, M. Algorithmic manipulation and information science: Media theories and cognitive warfare in strategic communication. *Eur. J. Commun. Media Stud.* **2025**, *4*, 1–11. [CrossRef]
76. Dietvorst, B.J.; Simmons, J.P.; Massey, C. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *J. Exp. Psychol. Gen.* **2015**, *144*, 114–126. [CrossRef]
77. Cadario, R.; Longoni, C.; Morewedge, C.K. Understanding, explaining, and utilizing medical artificial intelligence. *Nat. Hum. Behav.* **2021**, *5*, 1636–1642. [CrossRef]
78. Logg, J.M.; Minson, J.A.; Moore, D.A. Algorithm appreciation: People prefer algorithmic to human judgment. *Organ. Behav. Hum. Decis. Process.* **2019**, *151*, 90–103. [CrossRef]
79. Miller, T. Explanation in artificial intelligence: Insights from the social sciences. *Artif. Intell.* **2019**, *267*, 1–38. [CrossRef]

80. O'Donovan, J.; Smyth, B. Trust in recommender systems. In Proceedings of the 10th International Conference on Intelligent User Interfaces (IUI '05), San Diego, CA, USA, 10–13 January 2005; pp. 167–174. [[CrossRef](#)]
81. Ngoc, T.N.; Dung, M.V.; Pejić-Bach, M. Generation Z job seekers' expectations and their job pursuit intention: Evidence from transition and emerging economy. *Int. J. Eng. Bus. Manag.* **2022**, *14*, 18479790221112548. [[CrossRef](#)]
82. Beane, M. *The Skill Code: How to Save Human Ability in an Age of Intelligent Machines*; HarperCollins: New York, NY, USA, 2024.
83. Acquisti, A. The economics of privacy at a crossroads. In *The Economics of Privacy*; Goldfarb, A., Tucker, C.E., Eds.; University of Chicago Press: Chicago, IL, USA, 2023. Available online: <https://www.nber.org/books-and-chapters/economics-privacy/economics-privacy-crossroads> (accessed on 28 March 2026).
84. Jerman, A.; Pejić-Bach, M.; Aleksić, A. Transformation towards smart factory system: Examining new job profiles and competencies. *Syst. Res. Behav. Sci.* **2019**, *37*, 388–402. [[CrossRef](#)]
85. Lyon, D. *The Culture of Surveillance: Watching as a Way of Life*; Wiley: Hoboken, NJ, USA, 2018.
86. Nissenbaum, H. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*; Stanford University Press: Stanford, CA, USA, 2010.
87. Paul, J.; Lim, W.M.; O'Casey, A.; Hao, A.W.; Bresciani, S. Scientific procedures and rationales for systematic literature reviews (SPAR-4-SLR). *Int. J. Consum. Stud.* **2021**, *45*, O1–O16. [[CrossRef](#)]
88. Gan, Y.-N.; Li, D.-D.; Robinson, N.; Liu, J.-P. Practical guidance on bibliometric analysis and mapping knowledge domains methodology—A summary. *Eur. J. Integr. Med.* **2022**, *56*, 102203. [[CrossRef](#)]
89. Öztürk, O.; Kocaman, R.; Kanbach, D.K. How to design bibliometric research: An overview and a framework proposal. *Rev. Manag. Sci.* **2024**, *18*, 3333–3361. [[CrossRef](#)]
90. Tomaszewski, R.; Zhang, X.; Zuo, A.; Rafols, J. Visibility, impact, and applications of bibliometric and science mapping software tools: A review. *J. Informetr.* **2023**, *17*, 4007–4028. [[CrossRef](#)]
91. van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [[CrossRef](#)]
92. National Institute of Standards and Technology. *AI Risk Management Framework (AI RMF 1.0)*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2022. Available online: <https://www.nist.gov/itl/ai-risk-management-framework> (accessed on 4 April 2026).
93. von Krogh, G. Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing. *Acad. Manag. Discov.* **2018**, *4*, 404–409. [[CrossRef](#)]
94. Collins, C.; Dennehy, D.; Conboy, K.; Mikalef, P. Artificial intelligence in information systems research: A systematic literature review and research agenda. *Int. J. Inf. Manag.* **2021**, *60*, 102383. [[CrossRef](#)]
95. Shrestha, Y.R.; Krishna, V.; von Krogh, G. Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. *J. Bus. Res.* **2021**, *123*, 588–603. [[CrossRef](#)]
96. Braganza, A.; Chen, W.; Canhoto, A.; Sap, S. The impact of AI adoption on psychological contracts, job engagement and employee trust. *J. Bus. Res.* **2021**, *131*, 485–494. [[CrossRef](#)]
97. Jarrahi, M.H. AI's dual capacities for automating and informing work. *Bus. Inf. Rev.* **2019**, *36*, 178–187. [[CrossRef](#)]
98. Clifton, J.; Glasmeier, A.; Gray, M. When machines think for us. *Camb. J. Reg. Econ. Soc.* **2020**, *13*, 3–23. [[CrossRef](#)]
99. Pettersen, L. Why artificial intelligence will not outsmart complex knowledge work. *Work. Employ. Soc.* **2019**, *33*, 1058–1067. [[CrossRef](#)]
100. Bailey, D.E.; Faraj, S.; Hinds, P.J.; Leonardi, P.M.; von Krogh, G. We are all theorists of technology now. *Organ. Sci.* **2022**, *33*, 1–18. [[CrossRef](#)]
101. O'Neil, C. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*; Crown Publishing: New York, NY, USA, 2016.
102. Bessen, J. AI and Jobs: The Role of Demand. NBER Working Paper Series, No. 24235. 2018. Available online: <https://www.nber.org/papers/w24235> (accessed on 22 March 2026).
103. Brynjolfsson, E.; Rock, D.; Syverson, C. The productivity J-curve: How intangibles complement general purpose technologies. *Am. Econ. J. Macroecon.* **2021**, *13*, 333–372. [[CrossRef](#)]
104. Möhlmann, K.; Zalmanson, L. Hands on the wheel: Navigating algorithmic management and Uber drivers' autonomy. *Commun. ACM* **2017**, *60*, 62–67.
105. Zammuto, R.F.; Griffith, T.L.; Majchrzak, A.; Dougherty, D.J.; Faraj, S. Information technology and the changing fabric of organization. *Organ. Sci.* **2007**, *18*, 749–762. [[CrossRef](#)]
106. Poloz, S.S. Technological progress and monetary policy: Managing the fourth industrial revolution. *J. Int. Money Financ.* **2021**, *114*, 102373. [[CrossRef](#)]
107. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.M. How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* **2021**, *133*, 285–296. [[CrossRef](#)]
108. Kadlecová, L. *Cyber Sovereignty: The Future of Governance in Cyberspace*; Stanford University Press: Stanford, CA, USA, 2024.

109. Kaye, D. *Speech Police: The Global Struggle to Govern the Internet*; Columbia Global Reports: New York, NY, USA, 2019.
110. Mudambi, S.M.; Schuff, D. What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Q.* **2010**, *34*, 185–200. [[CrossRef](#)]
111. Narayanan, A.; Shmatikov, V. Robust de-anonymization of large sparse datasets. In Proceedings of the IEEE Symposium Security and Privacy, Oakland, CA, USA, 18–21 May 2008; pp. 111–125. [[CrossRef](#)]
112. Pejić-Bach, M.; Bertoneclj, T.; Meško, M.; Krstić, Ž. Text mining of Industry 4.0 job advertisements. *Int. J. Inf. Manag.* **2020**, *50*, 416–431. [[CrossRef](#)]
113. Pew Research Center. Views of Data Privacy, Risks, Personal Data, and Digital Privacy Laws. 2023. Available online: <https://www.pewresearch.org/internet/2023/10/18/views-of-data-privacy-risks-personal-data-and-digital-privacy-laws/> (accessed on 10 April 2026).
114. Pitke, M.; Joshi, V.; Baral, V.K.; Dwyer, R. Retailing and e-commerce riding on technology: Augmented reality and virtual reality. In *Augmented and Virtual Reality in Industry 5.0*; De Gruyter: Berlin, Germany, 2023; pp. 127–146.
115. Potamos, G.; Stavrou, E.; Stavrou, S. Enhancing maritime cybersecurity through operational technology sensor data fusion: A comprehensive survey and analysis. *Sensors* **2024**, *24*, 3458. [[CrossRef](#)]
116. Pritchard, A. Statistical bibliography or bibliometrics? *J. Doc.* **1969**, *25*, 348–349.
117. Jannach, D.; Manzoor, A.; Cai, W.; Chen, L. A survey on conversational recommender systems. *ACM Comput. Surv.* **2021**, *54*, 105. [[CrossRef](#)]
118. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144. [[CrossRef](#)]
119. Ricci, F.; Rokach, L.; Shapira, B. (Eds.) *Recommender Systems Handbook*, 2nd ed.; Springer: Cham, Switzerland, 2015. [[CrossRef](#)]
120. Roblek, V.; Meško, M.; Pejić-Bach, M.; Thorpe, O.; Šprajc, P. The interaction between Internet, sustainable development, and emergence of Society 5.0. *Data* **2020**, *5*, 80. [[CrossRef](#)]
121. Daza, A.; Rueda, N.D.G.; Quiñones, M.E.C. Sentiment analysis on e-commerce product reviews using machine learning and deep learning algorithms: A bibliometric analysis, systematic literature review, challenges and future works. *Int. J. Inf. Manag. Data Insights* **2024**, *4*, 100267. [[CrossRef](#)]
122. Khorshidi, M.S.; Merigó, J.M.; Beydoun, G. Half a century of *Information Processing & Management*: A bibliometric retrospective. *Inf. Process. Manag.* **2025**, *62*, 104238. [[CrossRef](#)]
123. Chen, W.; Li, A.; Sun, Y. Algorithmic management in the workplace: A systematic review and topic modeling integration using BERTopic. *Int. J. Inf. Manag.* **2026**, *86*, 102994. [[CrossRef](#)]
124. Munoko, I.; Brown-Liburud, H.L.; Vasarhelyi, M. The ethical implications of using artificial intelligence in auditing. *J. Bus. Ethics* **2020**, *167*, 209–234. [[CrossRef](#)]
125. Ferrara, E. Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci* **2024**, *6*, 3. [[CrossRef](#)]
126. Ge, Y.; Liu, S.; Fu, Z.; Tan, J.; Li, Z.; Xu, S.; Li, Y.; Xian, Y.; Zhang, Y. A survey on trustworthy recommender systems. *ACM Trans. Recomm. Syst.* **2024**, *3*, 1–68. [[CrossRef](#)]
127. Aria, M.; Cuccurullo, C. Bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975. [[CrossRef](#)]
128. Asaithambi, S.; Ravi, L.; Devarajan, M.; Almazyad, A.S.; Xiong, G.; Mohamed, A.W. Enhancing enterprises' trust mechanism through integrating blockchain technology into an e-commerce platform for SMEs. *Egypt. Inform. J.* **2024**, *25*, 100444. [[CrossRef](#)]
129. Zuboff, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*; PublicAffairs: New York, NY, USA, 2019.
130. Gombar, M.; Boban, M. Transferring marketing experiences from the military sector into the public sector: A conceptual framework. *Interdiscip. Descr. Complex Syst.* **2025**, *23*, 508–529. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.