**Published:** 05 March 2026
**Online AM:** 17 December 2025
**Accepted:** 19 November 2025
**Revised:** 19 October 2025
**Received:** 04 September 2024



SURVEY

# Overcoming Hazards of E-commerce Recommender Systems for Social Good

**ERANJANA KATHRIARACHCHI**, Massey University Auckland, Auckland, AUK, New Zealand

**SHAFIQ ALAM**, Massey University Auckland, Auckland, AUK, New Zealand

**SALMAN RASHID**, Massey University Auckland, Auckland, AUK, New Zealand

# Overcoming Hazards of E-commerce Recommender Systems for Social Good

ERANJANA KATHRIARACHCHI, School of Management and Marketing, Massey University Business School, Albany, New Zealand
SHAFIQ ALAM, School of Management and Marketing, Massey University Business School, Albany, New Zealand
SALMAN RASHID, School of Management and Marketing, Massey University Business School, Albany, New Zealand

Recommender systems used on e-commerce websites (e-commerce RS) are crucial to enhance the online shopping experiences of customers and improve business performance. Despite ongoing enhancements in their effectiveness, e-commerce RS face challenges that could harm customers, businesses, and society. Addressing these challenges effectively is crucial to protecting the interests of stakeholders relying on e-commerce RS. Proposed strategies include developing technical solutions, increasing customer awareness, and enacting relevant laws and regulations. However, current research has not yet thoroughly examined these solutions in a unified manner. This review fills that gap by providing a detailed overview of the hazards associated with e-commerce product recommendations and the measures implemented to mitigate these hazards. The paper assesses these solutions within the context of existing research, emphasizing their implications and suggesting future directions to improve the safety and efficacy of e-commerce RS.

CCS Concepts: • **Information systems** → *Recommender systems;*

Additional Key Words and Phrases: Information systems, information retrieval, retrieval tasks and goals, recommender systems, product recommendations, risks and hazards, solutions

**ACM Reference Format:**
Eranjana Kathriarachchi, Shafiq Alam, and Salman Rashid. 2026. Overcoming Hazards of E-commerce Recommender Systems for Social Good. *ACM Trans. Recomm. Syst.* 4, 3, Article 42 (March 2026), 34 pages. https://doi.org/10.1145/3785355

## 1 Introduction

**Recommender Systems** (**RS**) are now being used in various domains such as e-commerce, entertainment, healthcare, education, human resource management, online dating, and mobile applications [34, 53, 56, 66, 83, 102]. Algorithms enable RS to predict users' preferences and tailor platform offers accordingly. In e-commerce (hereinafter called e-commerce RS), they play a crucial role by helping customers navigate complex product information to find the most suitable options

Authors' Contact Information: Eranjana Kathriarachchi (corresponding author), School of Management and Marketing, Massey University Business School, Albany, New Zealand; e-mail: e.kathriarachchi@massey.ac.nz; Shafiq Alam, School of Management and Marketing, Massey University Business School, Albany, New Zealand; e-mail: s.alam1@massey.ac.nz; Salman Rashid, School of Management and Marketing, Massey University Business School, Albany, New Zealand; e-mail: m.s.rashid@massey.ac.nz.


ACM 2770-6699/2026/03-ART42
https://doi.org/10.1145/3785355

[24, 58, 87, 96]. Leading e-commerce platforms like Amazon, Alibaba, eBay, and Flipkart highlight the importance of RS in the business world [41, 46, 72, 101]. These e-commerce RS utilize both implicit data (such as users' online behavior) and explicit data (such as reviews and ratings) to build user profiles [23]. These profiles underpin personalized product recommendations to customers. Advances in RS technology have significantly boosted the growth of the e-commerce industry recently, resulting in higher revenues for businesses and better experiences for customers [13, 60].

Despite their contribution to the success of the e-commerce sector, RS suffer from various issues that may affect the welfare of users who rely on such systems. Scholars thus far have examined these issues from either an ethics perspective [13, 55, 64, 90] or a risk perspective [33, 40]. Researchers who adopt an ethical perspective have been discussing issues such as inappropriate content, privacy, autonomy, personal identity, opacity, fairness, and social effects. From a risk perspective, researchers have identified poor product decisions, negative user experience, and privacy breaches as areas of concern for customers. The ethical viewpoint is based on the premise that actions have consequences. This suggests that the ethical challenges of RS involve "any aspect of a RS that could negatively impact the utility of its stakeholders or risk imposing such negative impacts" [55]. For example, exposing a user to inappropriate content might result in some users adopting potentially damaging behaviors or consumption patterns. The risk perspective focuses mainly on the potential harmful effects of inferior recommendations. For example, poor product choices made by a customer may have direct financial consequences or result in a waste of time and resources. In addition, biased recommendations might lead to reduced customer trust, impacting the business performance [40]. As these explanations suggest, the ethical challenges and risks are often intertwined and share the common feature of causing harm or disadvantage to RS users.

Numerous researchers have emphasized the necessity of technology-based solutions, increasing customer awareness, and robust laws and regulations to address the challenges associated with e-commerce RS [13, 20, 72] because a combination of solutions is often more effective in mitigating the hazards associated with RS [28]. In response, a range of solutions has emerged across research, industry, and regulation domains. While several studies have investigated these solutions individually, focusing on specific technologies, domains, or particular risk areas (further elaborated in Section 2), no comprehensive review has been conducted that systematically examines the full spectrum of available responses to the risks posed by e-commerce RS. This lack of an integrated review of hazards (i.e., the entity or event that poses a negative consequence) and solutions limits our understanding of how those diverse solutions interact and complement one another in practice. To fill this gap, this review addresses the following research question: What solutions have been proposed to mitigate the hazards associated with e-commerce RS, and how can these be categorized into technological, customer, and regulatory dimensions to form a more effective response?

By synthesizing existing work across three key areas—technological solutions, customer awareness, and legal and regulatory measures—this review aims to provide a holistic view of the response landscape. In doing so, it offers valuable insights for researchers and practitioners aiming to enhance algorithmic fairness, stakeholder protection, and regulatory compliance in the context of e-commerce RS. The remainder of our paper is structured as follows. Section 2 provides an overview of related work to identify challenges associated with RS and potential solutions. Section 3 provides the necessary background on e-commerce RS hazards. Formal mathematical definitions for hazards, along with computable metrics, are presented in Sections 4. The methodology adopted for the scoping review is provided in Section 5. Section 6 further elaborates on the e-commerce RS hazards and presents the three categories of solutions available to mitigate these hazards. Section 7 offers a discussion based on our review. Implications and future research directions stemming from the review are presented in Section 8. Finally, Section 9 presents the conclusion of the study.

## 2 Related Work

In this section, we provide an overview of research studies examining the challenges related to RS and possible solutions to overcome them. We categorize these studies into three groups: technology-focused reviews, domain/application-specific reviews, and risk area-focused reviews.

### 2.1 Technology-focused reviews

Researchers investigating specific RS technologies have examined the challenges inherent in various techniques and proposed corresponding solutions to mitigate these challenges. Deep learning, a state-of-the-art machine learning approach, has been particularly influential in improving the quality of recommendations. In a comprehensive review [8] on deep learning-based RS, key challenges such as accuracy, data sparsity, the cold-start problem, and scalability were identified. The study provides detailed technical solutions to address these issues in the context of deep learning applications. Similarly, another study [30] has highlighted critical challenges in applying **Graph Neural Networks** (**GNNs**) in RS, including the use of deep GNNs, the dynamics of GNNs, and issues related to efficiency and scalability in large-scale GNN implementations. The technical solutions proposed in these studies aimed to enhance the effective utilization of specific technologies, thereby minimizing the negative impact.

### 2.2 Domain/application area-focused reviews

We identified several reviews that discuss challenges prevalent in specific domains of RS. In a study on music RS [78] cold start problem, automatic playlist continuation, and evaluating RS have been identified as challenges. With regard to news, the timeliness of reporting, user modeling, and quality control of the news content have been identified as challenges [75]. In the e-commerce domain, multiple studies have focused on identifying the challenges associated with e-commerce RS. A review study [5] has identified limitations related to big data problems and scalability as significant challenges associated with e-commerce RS. In another study [74] a range of issues, such as sparsity, scalability, privacy breaches, cold start problem, synonymy, diversity, and latency, have been discussed as challenges. When analyzing these reviews, it is evident that some challenges are highly domain-specific (e.g., timeliness of news in news RS), while others, such as the cold-start problem, extend beyond these specific domains examined in these studies and have broader applicability.

### 2.3 Risk area-focused reviews

Another approach taken by researchers to examine challenges associated with RS is to focus on a specific type of risk. One example of this is a review [28] on privacy aspects of RS, which identified three categories of solutions, namely (1) architecture, platforms, and standards, (2) algorithmic techniques, and (3) legislations, policies, and regulations, as measures to counter the negative impact or privacy breaches. In another similar survey [12] Seven types of biases (i.e., selection bias, exposure bias, conformity bias, position bias, inductive bias, popularity bias, and unfairness) associated with RS have been examined, along with methods to mitigate their effects. In this study, the researchers observed that various biases usually co-occur in real-world situations and suggested developing a general debiasing framework to manage the mixture of biases. Although several reviews on the e-commerce domain are presented here, they mainly focus on the technical solutions available to manage the challenges. Conversely, our review examines three different types of solutions that can work in tandem to manage the hazards associated with e-commerce RS. In addition, we base this review on hazards, as events with the potential to cause risk. This approach will be insightful for researchers interested in understanding the risk potential of e-commerce RS as a hazard domain.

## 3 Background

In this section, we first provide some background on e-commerce RS, followed by a brief overview of the hazards associated with e-commerce RS.

### 3.1 E-commerce RS

E-commerce RS are defined as "a web-based technology that explicitly or implicitly collects a consumer's preferences and recommends tailored e-vendors' products or services accordingly" [52]. This definition underscores the dual purpose of RS in supporting both recommendation providers (i.e., vendors or platforms) and recommendation receivers (i.e., customers). Consequently, e-commerce RS operate as multi-stakeholder systems [1, 40] in which customers, companies, and society are interconnected through algorithmic interactions. For e-commerce companies, RS have become essential tools to personalize digital experiences, increase conversion rates, and build customer loyalty, thereby improving overall business performance. From the customer perspective, RS help navigate information overload, identify relevant products, and make purchasing decisions more efficiently [13, 60].

Despite their advantages, e-commerce RS present several vulnerabilities that generate notable economic and social challenges, including unfair treatment of customers and producers, privacy breaches, and negative user experiences [31], ultimately eroding customer trust, reducing engagement, and diminishing revenue [40]. Given these wide-ranging implications, society is increasingly recognized as a critical stakeholder in the design and governance of RS [7, 40, 55]. For instance, recommendations can have an impact on a community of customers such as when it reinforces behavioral patterns in their collection behavior, leading to phenomena like echo chambers or filter bubbles. Against this backdrop, the next section of the review examines the risks and hazards associated with e-commerce RS.

### 3.2 Risks and Hazards of e-commerce Recommender Systems

The concept of risk has been widely discussed in various academic disciplines, such as consumer research, economics, medicine, and technology. A risk is defined as "the possibility of a negative outcome occurring as a result of exposure to a hazard" [93]. On the other hand, a hazard is defined as "the entity that poses a particular likelihood of negative consequences" [93]. These definitions suggest that hazards are specific events that can potentially cause risk to individuals. In the e-commerce RS domain, there are several studies [29, 33] which have adopted a hazard-based approach to explore risk perceptions among customers. In these studies, researchers first identified a series of hazards associated with the entire e-commerce lifecycle, which were then used to explore customer perceptions of the resulting risks. For instance, in the empirical study [33] these hazards are referred to as "unwanted events", which include certain events related to personalized recommendations provided by e-commerce RS. They have analyzed these hazards to conceptualize three types of risks, namely: Failure to gain product benefit risk, Functionality inefficiency risk, and Information misuse risk. This emphasizes the link between hazards (i.e., entities or events) and risk (i.e., the possibility of exposure to a hazard resulting in a negative outcome).

A more recent study [40] has proposed four types of risks associated with e-commerce RS, namely: Poor decision/choice dissatisfaction risk, Bad user experience/decision difficulty risk, Biased information state risk, and Privacy risk. For instance, a poorly designed RS can lead customers to make product decisions they will later regret, resulting in negative customer experiences. Poor product decisions are identified as a risk because they could lead to financial losses and waste of time for customers. The negative user experience is considered a risk because the difficulty in finding the exact information could lead customers to be stuck in an information bubble, resulting in a limited perspective. Other examples of risks include the potential for biased information to spread

Table 1. Risk Classifications on E-Commerce Recommendations

| **Risks [40]** | **Risks [33]** | **Explanation** |
|---|---|---|
| Failure to gain product benefit risk | Poor decision or choice dissatisfaction | Biased recommendations leading to inferior product decisions |
| Functionality inefficiency risk | Bad user experience / Decision difficulty risk | Poorly designed RS and improper recommendations leading to a poor user experience |
| Information misuse risk | Information misuse risk / Privacy | Unsolicited collection and use of confidential information |

community-related issues and the misuse of customer information, which can lead to privacy risks. When comparing the risk classifications proposed in these two studies, it is evident that, except for the biased information state risk [40], the other three types of risks [33] are similar to each other. These two risk classifications are presented in Table 1.

These studies have provided essential directions and insights into understanding the hazards and risks associated with e-commerce RS. However, they are limited in identifying a more comprehensive set of hazards related to e-commerce RS. For example, these studies do not identify hazards such as shilling/profile injection attacks and fake reviews, which can lead to suboptimal product decisions and perpetuate other decision biases. This lack of a comprehensive understanding of e-commerce RS hazards hinders accurate assessment of their risk potential. A recent scoping review [44] attempts to identify a more exhaustive list of hazards related to e-commerce RS. The hazards identified in this study are presented in Table 2.

*3.2.1 Biased recommendations.* E-commerce RS can provide various forms of bias, leading to suboptimal or manipulative suggestions that misalign with user preferences [89]. One prominent form is popularity bias, where RS favor widely-purchased or highly-rated products over others, potentially marginalizing niche or diverse options [13, 37, 46, 49, 61, 82, 91, 103]. This is also referred to as concentration bias [18, 26]. E-commerce platforms sometimes promote either highly profitable or less popular products for strategic reasons, such as profit maximization or inventory control [2, 24, 62, 87, 89, 94]. Additionally, exposure bias can occur when recommended systems sort products to favour those aligned with business interests, such as sponsored or high-priced items, regardless of user relevance [89].

These practices may lead to user distrust, negative experiences, and homogenization of choices, undermining the potential for better customer-product matches [24, 26, 62, 103]. Customers often seek novelty, diversity, and serendipity in recommendations, not just popular or similar items [23, 26, 32, 36, 46]. A key concern is that biased recommendations can mislead less informed users, who may trust the mass personalized suggestions [95]. Furthermore, anchoring bias may arise if users rate products based on prior recommended system suggestions rather than an independent assessment [2, 60].

The cold start problem, where recommended systems lack data on new users, is often addressed by recommending popular items; however, this strategy may overlook user-preferred unpopular items [4, 79]. Recommended systems can also be biased through incomplete or selective information, negatively influencing customer perception and decision-making [45, 97]. Additionally, visual bias, such as emphasizing certain product features or misrepresenting products with stereotypical imagery, can lead to dissatisfaction or exclusion of niche users [69, 88].

Collectively, these biases present critical ethical, commercial, and technical challenges. They not only impact user satisfaction and trust but also have broader implications for market diversity, seller fairness, and the transparency of automated decision-making systems. Addressing such biases

Table 2. Hazards of E-Commerce RS

| Categories | Risk-generating events / Hazards |
|---|---|
| Biased recommendations | Recommending popular products (popularity bias), less frequently purchased products, and/or disregarding unpopular, new, or obscure products (long tail products).<br>Presenting selective or incomplete information or concealing product details.<br>Promoting either profitable products or unprofitable products.<br>Using biased or unbalanced user data to provide skewed product recommendations to first-time or lesser-known customers.<br>Generating biased system ratings that influence subsequent customer preference ratings (anchoring effect).<br>Lacking an element of surprise (i.e., serendipity) and/or diversity in recommendations.<br>Presenting visually biased product recommendations and/or using biased marketing cues to promote products.<br>Implementing biased policies on exposing search results to customers (exposure bias). |
| Malicious activities | Shilling/profile injection attacks (push and/or nuke attacks) by malicious users.<br>Dishonest ratings by non-malicious users.<br>Obtaining, tracking, storing, using, or disclosing sensitive customer information in an unauthorized or undesired manner.<br>Reusing product reviews (review plagiarism) for unwarranted purposes. |
| Customer biases/actions | Customers complying with the choices of other customers (conformity/social influence bias).<br>Customers providing biased product reviews are influenced by the order in which existing online reviews are displayed (sequential bias).<br>Customers' decision biases (irregularities in human decision making), such as the tendency to make choices influenced by certain emotional states and/or interests.<br>Customers contributing to data leakages.<br>Customers favor a product simply because it is ranked high (position bias) or because they are influenced by neighboring/related items affecting the click-through rate of a target item (neighboring bias).<br>Customers lack awareness of how algorithms generate product recommendations. |
| Incompetent systems | Incompetent RS with functional issues (i.e., failure to consider complex user requirements, information overload, etc.).<br>RS fails to capture changing customer preferences over time (user and item bias). |

requires holistic approaches, including algorithmic debiasing, fairness-aware evaluation metrics, increased transparency, and user-centered design.

*3.2.2 Malicious activities.* Malicious activities threaten the integrity of e-commerce RS in several ways. A significant class of threats is shilling or profile injection attacks, in which adversaries build fake user profiles to sway system output [14, 15, 57, 80, 98, 104]. Within this category, push attacks artificially boost certain products, whereas new catalogs suppress them [14, 15]. When attackers create numerous false accounts, the result is a Sybil attack, which amplifies both push and nuke effects [14]. Even large platforms such as Amazon have suffered from such manipulation [16, 100]. Beyond rating inflation or deflation, adversary seed dishonest reviews that misrepresent

product quality, thereby regrading recommendation quality and accuracy, eroding customer trust, and nudging shoppers towards unsuitable purchases [4, 10, 14, 39, 57, 98, 101, 103]. Related tactic is review plagiarism, where users recycle existing reviews to promote their products or agendas, or to attack competitors [17]. These manipulations collectively diminish the quality of recommendations [16, 80, 101], compromise accuracy, and undermine customer trust [3, 100].

The next hazard involves privacy violations as recommended systems harvest ever larger troves of user data, and the incidence of data leakage has risen [13, 73, 77]. Privacy refers to the users' control over what data they share [72] remains a persistent worry in e-commerce, where shoppers expect their purchases and ratings to stay confidential [9, 84]. Breaches can occur during data transmission through sales to third parties or due to weak physical security [22, 51, 54, 67]. The personal footprints recommend that systems collect purchase histories, search keywords, preference profiles, and let attackers hit sensitive attributes such as income, interest, gender, and sexual orientation, for political learning that violates privacy [51, 54, 72, 73]. This data can also fuel price discrimination [22, 67] or unsolicited marketing [99] generating additional risks and negative user experiences.

These challenges underscore the dual vulnerability of e-commerce RS: susceptibility to manipulation and the erosion of user privacy. Malicious activities, such as shilling attacks and fraudulent reviews, compromise the reliability of recommendations, skewing outcomes in favor of strategic manipulation rather than genuine user relevance. Simultaneously, aggressive data collection practices and the increasing number of privacy breaches threaten user autonomy and trust, rendering personal data vulnerable to exploitation. Together, these issues highlight the urgent need for robust security mechanisms, transparent data handling policies, and resilient algorithmic defenses to safeguard both the integrity of recommendation outputs and the privacy rights of users. Without such safeguards, the long-term effectiveness and credibility of e-commerce RS remain at serious risk.

*3.2.3 Customer biases and actions.* Customers themselves can be a source of risk in e-commerce RS, primarily due to individual-level biases and a lack of awareness. One such bias is social influence bias or conformity bias, where customers are swayed by the reviews, ratings, and decisions of others rather than relying on their preferences [4, 35, 91]. Customers are also influenced by emotional status or temporary interest, which may lead to suboptimal decisions. Other cognitive biases include position bias, where users favor products ranked highly, and neighboring bias, where customers prefer items placed near their intended target [37]. For example, position bias can be illustrated by a customer clicking on a product simply because it is displayed at the top of the results. These biases can lead to suboptimal product choices. Moreover, customer-generated data, such as ratings and reviews, serves as a feedback loop that influences future recommendations [2, 4, 24]. If such input is biased, it degrades the quality of future suggestions. For instance, sequential bias occurs when the order of existing reviews influences new users, perpetuating skewed perceptions and further contaminating the data [23]. This feedback loop of biased customer input can lead to a long-term decline in recommendation quality and user satisfaction [4, 35].

Customer actions also contribute to privacy risks. Often, users accept terms and conditions without understanding the implications of data privacy, which intentionally facilitates data misuse [72]. Frequent interactions with various e-commerce platforms amplify this risk, as more personal data is shared with each new site [99]. Compounding this issue is the fact that customers typically do not understand how recommendations are generated, which may lead them to trust or act on suggestions without informed judgment [13]. Transparency and recommendation logic could mitigate this issue and reduce the likelihood of poor purchase decisions stemming from opaque algorithmic outputs [55].

This section highlights the often-overlooked role of users themselves in shaping the vulnerabilities of e-commerce RS. Individual cognitive and behavioral biases, such as conformity, position,

and emotional influence, can lead customers to make suboptimal choices while simultaneously feeding distorted data back into the system. These biases create self-reinforcing feedback loops that compromise the objectivity and long-term effectiveness of recommendations. Furthermore, users' limited understanding of privacy implications and recommendation mechanisms exacerbates the risks. As customers interact with multiple platforms, their personal data becomes increasingly fragmented and exposed, often without informed consent. These factors emphasize the importance of user education, transparency in recommendation logic, and designing systems that are not only technically robust but also aware of human cognitive tendencies and limitations.

*3.2.4 Incompetent systems.* Incompetence in e-commerce RS can negatively impact the user experience. One key issue is the failure to capture complex user requirements. Customers often struggle to articulate their specific needs or preferences accurately when interacting with recommended systems, making it difficult for the system to provide truly personalized recommendations [87]. Another major problem is information overload, which arises from the sheer volume of products, similar alternatives, and excessive related data. This can overwhelm customers, causing them to reject the system's recommendations or even switch to competing platforms [45]. The third concern is the inability to adapt to changing user preferences over time, referred to as temporal changes, as customers' interests, emotions, or contextual factors shift, so do their preferences. Yet many collaborative filter-based RS still rely on static historical data [21].

This section highlights the limitations of e-commerce RS stemming from system-level inefficiencies, which can significantly compromise the user experience. A primary challenge is the inability of many systems to accurately capture nuanced or evolving user needs, particularly when users struggle to express their preferences clearly. Additionally, the overwhelming volume of product options and related information often leads to information overload, which can prompt user frustration or abandonment of the platform. Another critical issue is the failure to adapt to temporal changes in user behavior, where many systems rely heavily on static historical data and thus struggle to reflect dynamic shifts in user interests or contexts. These shortcomings reveal the need for more adaptive, context-aware, and user-sensitive recommendation models to ensure long-term relevance and user engagement.

Building upon the hazards identified in [44], this review moves beyond conceptual classification to demonstrate that the hazards outlined in Section 3 are observable and measurable within real-world e-commerce RS. Before presenting potential solutions (Section 6), it focuses on operationalizing these risks–showing how they can be systematically detected, quantified, and evaluated in practice. This operational focus is essential as it (1) enables practitioners to monitor hazard severity, (2) establishes baseline metrics for assessing mitigation strategies, and (3) strengthens algorithmic accountability and transparency in automated decision systems [13, 55].

## 4 Hazard Metrics: Formalization and Evaluation

In this section, we demonstrate how the hazards identified in Section 3 can be operationalized as computable metrics using standard recommendation system logs and user data. We present formal mathematical definitions for ten critical hazard categories, each expressed as a quantifiable metric that can be directly implemented in production environments. To illustrate practical application, we construct a synthetic e-commerce dataset and compute each hazard metric, bridging the gap between conceptual hazard identification and empirical detection. This enables systematic evaluation of both hazard presence and the effectiveness of mitigation strategies proposed in Section 6.

### 4.1 Synthetic Dataset Specification

To demonstrate the practical application of our hazard metrics, we construct a synthetic e-commerce recommendation dataset simulating a realistic online retail scenario. This dataset serves as a running

Table 3. Synthetic Dataset with Recommendations and Relevance Judgments

| User | Group | Top-3 Recommendations |
|---|---|---|
| $u_1$ | Male | $i_1$ (E) ✓, $i_2$ X, $i_3$ (E) ✓ |
| $u_2$ | Male | $i_4$ (E) ✓, $i_5$ (M,F) ✓, $i_9$ (N) X |
| $u_3$ | Male | $i_7$ (M,F) X, $i_8$ (E) X, $i_2$ X |
| $u_4$ | Female | $i_1$ (E) X, $i_2$ ✓, $i_{10}$ (N) X |
| $u_5$ | Female | $i_5$ (M,F) X, $i_6$ (E) X, $i_3$ (E) X |
| $u_6$ | Female | $i_7$ (M,F) ✓, $i_8$ (E) ✓, $i_1$ (E) ✓ |

✓ = relevant (1), X = not relevant (0). Flags: (E) = explained, (N) = new, (M) = misinformation, and (F) = fake-influenced.

example, illustrating how each hazard can be detected and quantified using standard production system data.

The dataset reflects common e-commerce characteristics while incorporating deliberate hazard instances from Section 3. It represents a small-scale platform with six users receiving personalized recommendations. These include demographic groups for fairness assessment, new products for cold-start evaluation, and items flagged as misinformation or fake-influenced for content integrity risks, and explained versus unexplained recommendations for transparency assessment.

The system contains $m = 6$ users in two demographic groups (Male: $\{u_1, u_2, u_3\}$, Female: $\{u_4, u_5, u_6\}$) and $n = 12$ products ($I = \{i_1, \ldots, i_{12}\}$). Four items are new ($N = \{i_9, i_{10}, i_{11}, i_{12}\}$), two contain misinformation ($M = \{i_5, i_7\}$), and two are fake-influenced ($\mathcal{F} = \{i_5, i_7\}$). Each user receives $k = 3$ ranked recommendations, generating 18 total recommendation slots.

**Ground truth and labels:** Each user-item pair has ground truth relevance ($\text{rel}(u, i) \in \{0, 1\}$) indicating genuine preference matches. Items carry metadata flags: (E) for explained recommendations, (N) for new items, (M) for misinformation, and (F) for fake-influenced ratings. Table 3 presents the complete recommendation matrix.

Our suggested dataset is intentionally small for pedagogical clarity while demonstrating all hazard detection mechanisms. Real applications involve thousands of users and items, but computational principles remain identical. Multiple concurrent hazard types reflect production reality, necessitating comprehensive monitoring rather than isolated metrics.

### 4.2 Notation and Metric Definitions

The following formulations use standard RS notation. Users comprise the set $U = \{u_1, \ldots, u_m\}$ with $|U| = m$, while items form $I = \{i_1, \ldots, i_n\}$ with $|I| = n$. For each user $u$, the system generates a ranked recommendation list $R_u = (r_{u1}, \ldots, r_{uk})$ of length $k$. Ground-truth relevance is denoted $\text{rel}(u, i) \in \{0, 1\}$, indicating whether item $i$ genuinely matches user $u$'s preferences.

For hazard detection, users may be partitioned into groups $G$ based on demographic attributes, with $U_g \subseteq U$ representing users in group $g$. Three item subsets require tracking: new items $N \subseteq I$ introduced during the evaluation period, items containing misinformation $M \subseteq I$, and items with fake-influenced ratings $\mathcal{F} \subseteq I$. The function $\text{count\_rec}(i) = |\{(u, j) : r_{uj} = i\}|$ counts how many times item $i$ appears across all recommendations, while $\mathbf{1}\{\cdot\}$ denotes the standard indicator function.

The total recommendation slots equal $m \cdot k$. Group-level metrics aggregate individual user metrics using arithmetic mean unless specified otherwise. The following subsections define ten hazard metrics corresponding to the taxonomy in Section 3, each demonstrating a formal definition, computation on our synthetic dataset, and interpretation of hazard severity. These metrics enable monitoring system safety and provide baselines for evaluating mitigation strategies in Section 6.

*4.2.1 Biased Product Recommendations.* This metric detects algorithmic bias by measuring relevance disparities between demographic groups. Following the hazard identified in Section 3, we assess whether the RS provides systematically different recommendation quality to different user groups.

**Formal definition:** We define per-user precision at $k$ as:

$$P@k(u) = \frac{1}{k} \sum_{j=1}^{k} \text{rel}(u, r_{uj}). \tag{1}$$

This measures the proportion of relevant items in user $u$'s top-$k$ recommendations. Group-level precision aggregates individual user precision scores within demographic group $g$:

$$P@k(g) = \frac{1}{|U_g|} \sum_{u \in U_g} P@k(u). \tag{2}$$

The bias ratio quantifies the relative disparity between two groups:

$$\text{Bias Ratio}(g_1, g_2) = \frac{P@k(g_1)}{P@k(g_2)}. \tag{3}$$

Values significantly different from 1 indicate bias. We also report the absolute difference $|P@k(g_1) - P@k(g_2)|$ to capture the magnitude of disparity.

**Computation on synthetic dataset:** Per-user precision at $k = 3$:

$$P@3(u_1) = \frac{2}{3} \approx 0.667, \quad P@3(u_2) = \frac{2}{3} \approx 0.667, \quad P@3(u_3) = \frac{0}{3} = 0.000$$
$$P@3(u_4) = \frac{1}{3} \approx 0.333, \quad P@3(u_5) = \frac{0}{3} = 0.000, \quad P@3(u_6) = \frac{3}{3} = 1.000.$$

Group precision:

$$P@3(\text{male}) = \frac{1}{3}(0.667 + 0.667 + 0.000) = \frac{4}{9} \approx 0.444$$
$$P@3(\text{female}) = \frac{1}{3}(0.333 + 0.000 + 1.000) = \frac{4}{9} \approx 0.444.$$

Bias ratio: $\text{BiasRatio}(\text{male}, \text{female}) = \frac{4/9}{4/9} = 1.000$

Both groups receive recommendations with identical average precision (0.444), yielding a bias ratio of exactly 1.000 and zero absolute difference. While individual users exhibit varying recommendation quality (ranging from 0.000 to 1.000), these variations balance equally across groups, indicating the system does not systematically favor one demographic over another. This metric would flag bias if, for example, male users consistently received higher-quality recommendations than female users.

*4.2.2 Cold-Start Problem.* This metric quantifies the system's ability to surface new items to users, addressing the hazard identified in Section 3, where new products may be systematically under-recommended.

**Formal definition:** New item coverage measures the proportion of new items receiving at least $t$ recommendations:

$$\text{Coverage}_{\text{new}}(t) = \frac{|\{i \in N \,:\, \text{count\_rec}(i) \geq t\}|}{|N|}, \tag{4}$$

where $t$ is the minimum recommendation threshold, typically set to $t = 1$ or $t = 2$.

**Computation on synthetic dataset:** Item recommendation counts yield $\text{count\_rec}(i_9) = 1$, $\text{count\_rec}(i_{10}) = 1$, $\text{count\_rec}(i_{11}) = 0$, $\text{count\_rec}(i_{12}) = 0$. New item coverage equals $\text{Coverage}_{\text{new}}(1) = \frac{2}{4} = 0.500$.

Half of the new items receive at least one recommendation, indicating moderate cold-start handling. Two new items ($i_9$ and $i_{10}$) successfully entered recommendation lists, while two others ($i_{11}$ and $i_{12}$) received no exposure. This 50% coverage suggests the system provides some visibility to new products but fails to comprehensively address the cold-start problem, potentially disadvantaging newer marketplace entrants.

*4.2.3 Privacy Breaches.* This metric measures vulnerability to adversarial inference attacks, where attackers attempt to infer sensitive user attributes from recommendation patterns, addressing the privacy hazard discussed in Section 3.

**Formal definition:** Let $A$ be an adversary attempting to infer user attribute $x_u$ from recommendation lists, with success probability $p_u = P_A(\hat{x}_u = x_u)$ for user $u$. Privacy risk is defined as:

$$\text{Privacy Risk} = \frac{1}{m} \sum_{u \in U} p_u. \tag{5}$$

In practice, $p_u$ is estimated by training attack models on recommendation data and measuring per-user prediction accuracy.

**Computation on synthetic dataset:** Using simulated adversarial success probabilities $p_{u_1} = 0.90$, $p_{u_2} = 0.80$, $p_{u_3} = 0.70$, $p_{u_4} = 0.60$, $p_{u_5} = 0.40$, $p_{u_6} = 0.30$, we obtain PrivacyRisk $= \frac{0.90+0.80+0.70+0.60+0.40+0.30}{6} = 0.617$.

The 61.7% average adversarial success rate indicates substantial privacy risk. An adversary can correctly infer sensitive user attributes (such as gender in our demographic groups) from recommendation patterns in over 60% of attempts. This demonstrates how recommendation lists can leak personal information even without explicit user data disclosure, motivating the privacy-preserving solutions discussed in Section 6.

*4.2.4 Fake Reviews and Ratings Impact.* This metric quantifies how fake content influences recommendation quality metrics, addressing the malicious activity hazard from Section 3.

**Formal definition:** For any quality metric $M(\cdot)$, let $M_{\text{all}}$ denote the metric computed on full data and $M_{\neg \mathcal{F}}$ the metric after removing recommendations driven by fake-influenced items. Fake content impact is:

$$\text{Fake Impact} = \frac{M_{\text{all}} - M_{\neg \mathcal{F}}}{M_{\text{all}}}. \tag{6}$$

Positive values indicate metric inflation due to fake reviews.

**Computation on synthetic dataset:** Global precision on full data yields $M_{\text{all}} = \frac{8}{18} = \frac{4}{9} \approx 0.444$. After removing four fake-influenced recommendation slots (items $i_5$ and $i_7$, of which two were relevant), precision becomes $M_{\neg \mathcal{F}} = \frac{6}{14} = \frac{3}{7} \approx 0.429$. Thus, FakeImpact $= \frac{0.444-0.429}{0.444} \approx 0.036$.

Fake reviews inflate the apparent precision by 3.6%. While this impact appears modest in our synthetic example, it demonstrates how manipulated ratings can artificially boost system quality metrics. In production systems with higher fake content prevalence, this inflation could substantially misrepresent true recommendation quality and mislead both users and platform operators.

*4.2.5 Biased Customer Feedback.* This metric measures systematic skew in observational feedback relative to ground truth, addressing the customer bias hazard from Section 3.

**Formal definition:** Let $P_{\text{obs}}(r)$ be the empirical distribution of observed ratings and $P_{\text{true}}(r)$ the unbiased distribution from surveys or ground-truth samples. Feedback skew is quantified using KL divergence:

$$\text{Feedback Skew} = D_{KL}(P_{\text{obs}} \| P_{\text{true}}) = \sum_r P_{\text{obs}}(r) \ln \frac{P_{\text{obs}}(r)}{P_{\text{true}}(r)}. \tag{7}$$

**Computation on synthetic dataset:** Using example distributions from the literature, $P_{\text{true}} = \{1 : 0.10, 2 : 0.10, 3 : 0.20, 4 : 0.30, 5 : 0.30\}$ and $P_{\text{obs}} = \{1 : 0.05, 2 : 0.05, 3 : 0.20, 4 : 0.40, 5 : 0.30\}$, we obtain FeedbackSkew $= 0.046$ nats.

Moderate observational feedback skew detected. The observed ratings distribution deviates from ground truth, with users under-reporting low ratings (1–2 stars) and over-reporting moderate-high ratings (4 stars). This skew can arise from social desirability bias or self-selection, where dissatisfied customers disengage rather than provide negative feedback. Such systematic distortions contaminate training data and propagate bias through future recommendations.

*4.2.6 Biased Decision Making.* This metric detects downstream decision disparities based on recommendation-derived metrics, extending the bias analysis beyond recommendation quality to consequential outcomes.

**Formal definition:** For decisions based on threshold $\tau$ applied to user metric $Q(u)$, group decision rate is:

$$\text{Decision Rate}(g) = \frac{1}{|U_g|} \sum_{u \in U_g} \mathbf{1}\{Q(u) \geq \tau\}. \tag{8}$$

Decision bias between groups is quantified as:

$$\text{Decision Bias}(g_1, g_2) = \frac{\text{Decision Rate}(g_1)}{\text{Decision Rate}(g_2)}. \tag{9}$$

**Computation on synthetic dataset:** Using decision threshold $\tau = 0.5$ on $P@3(u)$ yields decision indicators: $u_1 : 1, u_2 : 1, u_3 : 0, u_4 : 0, u_5 : 0, u_6 : 1$. Group decision rates are DecRate(male) $= \frac{2}{3} \approx 0.667$ and DecRate(female) $= \frac{1}{3} \approx 0.333$. Decision bias equals DecisionBias(male, female) $= \frac{2/3}{1/3} = 2.000$.

Males receive favorable decisions twice as often as females. While Section 4.2.1 showed no group-level precision bias, downstream decisions exhibit substantial disparity. This demonstrates how seemingly fair recommendations can produce biased outcomes when thresholds are applied, a critical concern for systems driving consequential decisions such as product promotions, personalized pricing, or service tier assignments.

*4.2.7 Incompetent Recommender Systems.* This metric combines multiple quality dimensions into an overall competence assessment, addressing the system inadequacy hazard from Section 3.

**Formal definition:** Item coverage measures the proportion of catalog items receiving recommendations:

$$\text{ItemCoverage} = \frac{|\{i \in I : \text{count\_rec}(i) \geq 1\}|}{|I|}. \tag{10}$$

Overall competence aggregates multiple quality metrics:

$$\text{Competence} = \alpha \cdot P@k_{\text{global}} + \beta \cdot \text{Item Coverage} + \gamma \cdot \text{Diversity}, \tag{11}$$

where $\alpha + \beta + \gamma = 1$ and the system is flagged as incompetent if Competence $< \theta$.

**Computation on synthetic dataset:** Global precision equals $P@k_{\text{global}} = \frac{4}{9} \approx 0.444$. Item coverage yields ItemCoverage $= \frac{10}{12} \approx 0.833$ (10 unique items recommended out of 12 total). Using weights $\alpha = 0.7, \beta = 0.3$, competence becomes Competence $= 0.7 \times 0.444 + 0.3 \times 0.833 = 0.561$.

The system competence of 0.561 exceeds the 0.5 threshold, indicating adequate but not exceptional performance. The system demonstrates moderate precision (44%) and strong catalog coverage (83%), suggesting it avoids extreme popularity bias. However, the modest precision indicates substantial room for improvement in matching users with relevant items. Production systems should set competence thresholds based on business requirements and user satisfaction benchmarks.

*4.2.8 Algorithmic Opacity.* This metric measures the availability of human-interpretable explanations for recommendations, addressing the transparency hazard from Section 3.

**Formal definition:** Let $e_{uj} \in \{0, 1\}$ indicate whether an explanation accompanies recommendation slot $(u, j)$. The transparency index is:

$$\text{TransparencyIndex} = \frac{1}{m \cdot k} \sum_{u} \sum_{j=1}^{k} e_{uj}. \tag{12}$$

Lower values indicate higher opacity hazard.

**Computation on synthetic dataset:** Nine out of 18 total recommendation slots have explanations available (see Table 3), yielding TransparencyIndex $= \frac{9}{18} = 0.500$.

The 50% explanation coverage indicates moderate algorithmic opacity. Half of the recommendations lack human-interpretable justifications, potentially reducing user trust and hindering informed decision-making. Users receiving unexplained recommendations (such as $u_3$ with items $i_7, i_2$, or $u_4$ with item $i_{10}$) cannot understand why these products were suggested. This opacity becomes particularly problematic when recommendations prove irrelevant or when users seek to understand system behavior.

*4.2.9 Spreading of Misinformation.* This metric quantifies user exposure to identified misinformation content, addressing the content integrity hazard from Section 3.

**Formal definition:** Misinformation exposure measures the proportion of recommendation slots containing flagged misinformation items:

$$\text{MisinfoExposure} = \frac{1}{m \cdot k} \sum_{u} \sum_{j=1}^{k} \mathbf{1}\{r_{uj} \in M\}. \tag{13}$$

**Computation on synthetic dataset:** Four out of 18 recommendation slots contain misinformation (items $i_5$ and $i_7$ each appear twice), yielding MisinfoExposure $= \frac{4}{18} = \frac{2}{9} \approx 0.222$.

Over 22% of recommendations contain misinformation, indicating substantial exposure to potentially harmful content. Items $i_5$ and $i_7$, despite being flagged as containing misinformation, appear in recommendations for users $u_2, u_3, u_5$, and $u_6$. This high exposure rate demonstrates how content integrity failures can propagate through recommendation systems, potentially misleading users and eroding platform credibility. The mitigation strategies in Section 6 address this through content moderation and quality filters.

*4.2.10 Unfair Treatment of Customers.* This metric measures group-level utility disparities relative to global performance, providing an aggregate assessment of fairness across demographic groups.

**Formal definition:** For utility metric $T(g)$ computed for group $g$, unfairness is quantified as the average absolute deviation from global performance:

$$\text{Unfairness} = \frac{1}{|G|} \sum_{g \in G} |T(g) - T_{\text{global}}|, \tag{14}$$

where $T_{\text{global}}$ is the metric computed over all users.

**Computation on synthetic dataset:** Using precision as the utility metric, group utilities equal $T(\text{male}) = 0.444$ and $T(\text{female}) = 0.444$, with global utility $T_{\text{global}} = 0.444$. Group deviations from global yield $|0.444 - 0.444| = 0$ for both groups, thus Unfairness $= \frac{0+0}{2} = 0.000$.

No group-level unfairness detected. Both demographic groups receive identical average recommendation quality, resulting in zero deviation from global performance. This aggregate fairness measure complements the bias ratio metric (Section 4.2.1) by assessing distributional equity across

Table 4. Hazard Metric Values Computed on the Synthetic Dataset

| Hazard Metric | Value | Interpretation |
|---|---|---|
| BiasRatio (male/female) | 1.000 | No detected relevance bias |
| $\text{Coverage}_{\text{new}}(1)$ | 0.500 | 50% new item coverage |
| PrivacyRisk | 0.617 | High adversarial inference risk |
| FakeImpact | 0.036 | 3.6% metric inflation from fakes |
| FeedbackSkew | 0.046 | Moderate observational bias |
| DecisionBias | 2.000 | 2× male advantage in decisions |
| Competence | 0.561 | Above-threshold system performance |
| TransparencyIndex | 0.500 | Moderate algorithmic opacity |
| MisinfoExposure | 0.222 | 22% misinformation exposure |
| Unfairness | 0.000 | No detected group unfairness |

all groups simultaneously. However, this metric masks within-group variation, as individual users within each group experienced widely varying recommendation quality (ranging from 0.000 to 1.000 precision).

### 4.3 Operationalizing Hazard Detection in Practice

Table 4 consolidates all hazard metric values computed on our synthetic dataset, providing a comprehensive view of the system's safety profile.

These formalized metrics provide a systematic approach to hazard detection in RS. All metrics can be computed using standard recommendation logs, user demographics, and content labels, making them practical for production deployment. Privacy risk estimation requires training adversarial models on held-out users to estimate inference success rates. Decision thresholds and competence weights should be set by domain experts and stakeholders based on business requirements and regulatory constraints. Metrics should be computed over sliding time windows to detect hazard emergence or mitigation effectiveness over time. Production deployments should include confidence intervals and significance testing to ensure statistical robustness.

This framework enables systematic hazard monitoring and provides quantitative targets for algorithmic auditing and regulatory compliance. Having demonstrated that the hazards identified in Section 3 can be operationalized as computable metrics, we now examine the solutions proposed to mitigate these risks. The following section details the methodology for our scoping review of technological, customer awareness, and regulatory interventions.

## 5 Methodology

We conducted a review of the literature to identify academic publications on minimizing hazards associated with e-commerce RS. The concepts and the corresponding search terms were used to develop a comprehensive search query. The search was conducted through the SCOPUS database and the EBSCOhost platform. Table 5 outlines the concepts and the respective search terms used.

Inclusion and exclusion criteria were used to select the publications for this review. Only studies examining/referring to hazards and possible solutions in the context of e-commerce RS were considered for this review. Only peer-reviewed studies (journal articles and conference papers) were included in this review, as they are deemed to have higher validity than other types of publications. No time limit was applied in searching for academic publications. Only English language publications were considered for this review. The **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)** guidelines were followed in identifying suitable publications for the review [85]. A total of 68 records were identified in the initial search.

Table 5. Concepts and Search Terms

| Concept | Search terms |
|---|---|
| RS | "recommend* system*" OR "recommend* agent*" |
| Context | e-commerce OR "e-business" OR "electronic commerce" OR "digital business" OR "online business" |
| Hazards and solutions | ("risk*" OR "hazard*") AND ("mitigation" OR "reduction" OR "management") |

Table 6. Selected Publications

| Hazard | Publications |
|---|---|
| Biased product recommendations | [36, 87, 94] |
| Cold-start problem | [50] |
| Privacy breaches | [6, 27, 52, 77, 99] |
| Incompetent RS | [32, 48, 87] |

The search results were then compared, and duplicates were removed. After that, 15 records were removed because they were conference proceedings and did not meet the criteria for peer-reviewed empirical studies. Fifty articles were subjected to full-text review. After this stage, thirty-four publications were excluded for the following reasons: they were in an irrelevant domain (not the e-commerce sector), had an irrelevant outcome (absence of hazards and/or solutions), or were of an irrelevant study type (i.e., reviews). Eleven articles were identified as suitable for this review. The data extracted included the hazards discussed and the proposed solutions.

The 11 publications identified through the review were categorized by hazard, as indicated in Table 6. To enhance the comprehensiveness and depth of the analysis, additional relevant literature was incorporated from the prior review that examined similar themes. This approach enabled a more robust synthesis of existing knowledge, ensuring alignment with established academic insights in the field. Through this exercise, research publications proposing solutions to mitigate the hazards associated with e-commerce RS were identified. In line with the research question, the following section presents these solutions categorized under technological, customer, and regulatory dimensions.

## 6 Hazard Solutions

The diverse range of previously introduced hazards necessitates identifying practical solutions to mitigate their impact. Specifically, this section examines how both customers and companies can optimize the benefits of e-commerce RS while mitigating the associated risks. We categorize these solutions into three key areas: technological solutions, increased customer awareness, and laws and regulations. Additionally, the discussion will further elaborate on the nature of these hazards with the proposed solutions. A summary of the hazards and their corresponding solutions is presented in Table 7.

### 6.1 Technological solutions

By technological solutions, we primarily refer to a wide range of solutions designed to address various issues associated with e-commerce RS. These include biased product recommendations, cold-start problem, privacy breaches, fake reviews and ratings, incompetent e-commerce RS, and unfair treatment of customers.

Table 7. Hazard and Solutions

| Hazards | Technological solutions | Increasing customer awareness | Laws and regulations |
|---|---|---|---|
| Biased product recommendations | X | | X |
| Cold-start problem | X | | |
| Privacy breaches | X | X | X |
| Fake reviews and ratings | X | | |
| Biased customer feedback | X | | |
| Biased decision-making | | X | |
| Incompetent RS | X | | |
| Algorithmic opacity | | X | X |
| Spreading of misinformation | | | X |
| Unfair treatment of customers | X | | X |

*6.1.1 Biased product recommendations.* E-commerce RS can present biased recommendations in multiple ways. Biased recommendations are identified as a hazard because they can lead customers to make inferior product choices, resulting in negative user experiences. Researchers have identified these diverse hazards and developed technological solutions to mitigate their impact on the recommendation process.

One common hazard associated with e-commerce RS is the bias toward popular products (e.g., RS disproportionately recommending a product to all customers due to its popularity). Some researchers have identified this as a fairness issue, where less popular brands have fewer opportunities to be seen by potential customers, thereby limiting their potential for achieving business success. Several techniques have been experimented with by researchers to address this issue. One such technique is the application of a learning-to-rank method with fairness-aware regularization to increase the proportion of medium-tail recommendations for non-popular items without compromising ranking performance [46]. Similarly, another fairness-aware technique is the **Fair Personalized Recommendation Algorithm** (**FPRA**), which can simultaneously achieve both accuracy and fairness in recommendation results [103]. Another technique aimed at the same outcome is the two-injection-based approach [18], which provides visibility to both long-tail items and users. The fuzzy query approach is capable of recommending less frequently purchased products to customers based on fuzzy conformance between features appearing in the vector of the ideal product from the customer's point of view and the features of existing products [87]. In a more recent study [71] researchers have introduced CP-FairRank, an optimization-based re-ranking algorithm that integrates fairness constraints from both the customer and business sides in a joint objective framework. These algorithms have recognized the two-sided nature of the RS functionality in the marketplace. When examining the solutions, it is evident that researchers have identified the need to establish fairness in the RS context by providing a balanced level of visibility to both popular and non-popular products without compromising the accuracy of recommendations. Such solutions will help mitigate the effect of popularity bias, providing customers with more diverse product recommendations and fair access to the market for small-scale producers whose products are not significantly popular.

Another type of biased recommendation discussed by researchers is the presentation of selective or incomplete information. This occurs when e-commerce companies deliberately provide selected information while neglecting negative comments or hiding negative terms to attract more customers and achieve higher profitability [97]. This could lead to an impaired user experience and gradually undermine customers' confidence in the recommendations [45, 97]. The "RecRisk" integrates

trust, heat equation, and modern portfolio theory to address issues of incomplete or misleading information [97]. This technique allows customers to choose trusted services that align with their preferences by constructing and solving a **trust-aware heat equation** (**TAHE**), which considers variations in personal preferences and trust-aware factors among customers. Promoting profitable products is another reason why recommendations can be biased. Providing warning messages or explanations on organic recommendations and sponsorship disclosures can help customers identify biases in product recommendations, helping them avoid this problem [89, 94].

It is becoming increasingly clear that achieving high recommendation accuracy alone does not ensure an adequate or satisfying experience for users of RS [65]. Hence, sometimes it's better to recommend a variety of related items rather than a narrowly targeted list of products [81]. Proponents of concepts such as diversity, novelty, and serendipity think that excessive focus on accuracy could lead to concentration bias [18], causing dissatisfaction among customers. Hence, the prudent approach is to balance quality factors such as diversity, novelty, and serendipity with prediction accuracy [65, 76]. Researchers propose several mechanisms to overcome this challenge. A user-based two-step recommendation with popularity normalization, designed to improve recommendation diversity and novelty, is one such technique [61]. This method analyzes customer behaviors embedded in rating data and provides recommendations that integrate user ratings, user similarity, and item popularity. The **Trigger and Triggered** (**TT**) model addresses concentration bias by generating meaningful product pairs based on sequential purchase behavior rather than raw frequency. It introduces a normalized tt-score that reduces the dominance of popular items by balancing the influence of each triggered product. The model further refines these pairs using expert input and linear programming to ensure diversity and marketing alignment. Consequently, it significantly reduces the Gini coefficient, indicating improved diversity in recommendations [18].

Another issue with biased recommendations is the presentation of visually biased recommendations or the use of biased marketing cues. Customers usually tend to click and view products based on the visual satisfaction of their expectations [70]. However, a real purchase does not always occur due to their dissatisfaction with other essential features such as brand, material, and price. When RS cannot identify the deviation between visually related interaction and actual preference, it might lead to a visually biased recommendation. One technological solution to minimize this hazard is the use of causal graphs, which identify and analyze the potential visual biases [70]. A causal graph is derived from the existing recommendation methods. In a causal graph, the visual feature of an item acts as a mediator, potentially introducing a spurious relationship between the customer and the item. The spurious relationship that could mislead the prediction of the customer's real preference is eliminated by intervention, and a counterfactual inference is developed over the mediator. In the same study, the researchers introduced a visually aware RS, named CasualRec, which can retain the supportive significance of visual information and remove visual bias. There are also instances when biased marketing cues are used in e-commerce RS. Marketing bias occurs when strategies, such as using a human model in a product image, are used in a biased way, influencing customer-product interactions [64]. At times, the marketing strategies of companies can introduce biases in modern RS, potentially harming business performance, customer experience, and the broader society. In a study [88] a framework has been developed to mitigate the effects of these potential marketing biases. Their framework has demonstrated the capability to provide increased exposure to underrepresented market segments without significantly affecting the accuracy of recommendations.

Exposure bias, which refers to biased practices in the way search results are displayed, is yet another problem affecting customers who rely on e-commerce product recommendations. To overcome the exposure bias, a novel recommendation framework called the InvPref has been proposed. It iteratively decomposes the invariant preferences and variant preferences from biased

observational user behaviors by heterogeneous environments corresponding to different types of latent biases [91]. Reviewing these different techniques shows that a substantial amount of research is being conducted to mitigate various types of biases in product recommendations. If these techniques are successfully implemented in commercial applications, they could provide safer and more beneficial online shopping experiences for customers, fostering enhanced trust and loyalty towards the businesses.

*6.1.2 Cold-start problem.* The cold-start problem occurs when RS lacks the necessary data to provide users with meaningful recommendations. This is common on platforms with fewer users and products, or in situations where a new customer lacks a significant product history. This becomes a hazard when recommendations provided with a lack of data lead to the risks of not meeting the needs of the customers, or their online shopping experience becomes poor [38, 79].

To address the cold-start problem, researchers have developed several solutions. Cross-domain RS, which transfers knowledge from domains with dense ratings to other domains with sparse ratings [4, 50] is one such solution. **Trust-Aware Spatial-Temporal Activity-based Denoising Autoencoder** (**TSTDAE**), as another technique, not only considers the time at finer granularity levels but also considers a customer's location, trust level, and sentiment analysis when computing recommendations [4]. Another solution to overcome the cold-start problem is RS based on user coverage maximization: Max-coverage and Category Exploration [79]. These techniques aim to explore customer coverage to diversify the recommended items and attract more first-time customers. In addition, the coupled regularization approach consists of two latent factor models (C-HMF and S-HMF) [38]. These two models can produce more qualitative predictions for both tail users and tail items. Similarly, two injection-based approaches [82] are capable of improving the quality of recommendations to the users at the tail of the distribution. Different techniques have been developed that can provide more prominence to less well-known users and products within e-commerce RS. This is crucial in giving fair treatment to diverse types of customers who do not belong to the mainstream customer groups.

*6.1.3 Privacy breaches.* Privacy is the ability of an individual to determine what information can be shared and employ access control [72]. The increasing number of privacy breaches has become a significant concern due to the large volumes of data used to personalize product recommendations [73, 77]. Privacy breaches can occur when transmitting data within the RS, selling it to third parties, failing to provide the required physical security, or due to the actions of other users [22, 28, 51, 54]. Privacy concerns could result in a negative user experience for customers. Due to this, privacy preservation has been a key interest among e-commerce RS researchers who have developed a wide range of techniques to achieve this aim.

One key method to mitigate privacy risks is to obscure the link between users and the data collected [13]. Homomorphic encryption is one technique to obscure the private rating information of the customers from the e-commerce company [22]. This method enables the company to generate recommendations using encrypted customer ratings while protecting the company's item-item similarities. Another technique is the **privacy-preserving method** (**PRS**) for e-commerce RS, which is also capable of obscuring the link between customers and the data collected [43]. PRS provides customers with solutions to their privacy concerns and reduces error rates. PRS first uses an anonymous method to convert secondary data without user identification.

Researchers have developed some techniques that give customers autonomy to safeguard their privacy. With a local differential privacy mechanism for **matrix factorization** (**MF**)-based RS, customers can perturb their ratings locally on their devices and send the perturbed ratings to the e-commerce companies [42]. Blockchain-supported secure multiparty computation can also be used to strengthen the cryptographic security of customer data [27]. Using this method, a potential

customer can allow a company to apply a recommendation algorithm without disclosing their personal data.

E-commerce RS has a big data application, that is, prone to inference attacks. One technique that can potentially address this issue is the data lake-based modernistic privacy preservation technique, which helps to handle privacy preservation in unstructured data [72]. At the same time, MF has contributed to the success of RS. However, designing MF with privacy preservation could lead to issues such as error accumulation in multiple iterations of MF, the introduction of unnecessary noise, and difficulties in sensitivity analysis. Using a **vector perturbation-based differentially private matrix factorization** (**VP-DPMF**) to overcome these issues [73]. VP-DPMF prevents error accumulation by perturbing the objective function of MF rather than its factorization process or results.

Using large volumes of customer data increases the possibility of attackers getting more sensitive information. To counter this, a **short-term dynamic recommendation model based on local differential privacy** (**SDRM-LDP**) is introduced [51]. This model utilizes a limited amount of user information to construct short-term user preference behaviors. In addition, a three-tier RS architecture has been proposed to manage this situation [99]. These three tiers include a customer tier, a third-party server tier, and the vendor e-store tier. This new structure reduces the risk of privacy disclosure and also improves the quality of recommendations to new users.

*6.1.4 Fake reviews and ratings.* Fake reviews and ratings have become a significant issue that compromises the quality of recommendations provided to customers. The reviews and ratings become input in the recommendation process. Therefore, fake reviews and ratings become hazardous when they influence the quality of the recommendations, leading to inferior product decisions and negative customer experiences. Fake ratings can enter the recommendation process through shilling or profile injection attacks (e.g., a seller artificially inflating ratings for a product by posting fake reviews). Profile injection attacks can take the form of push attacks or nuke attacks. Push attacks involve giving higher ratings, while nuke attacks involve giving lower ratings maliciously [15].

There is a myriad of technological solutions proposed to curtail shilling and profile injection attacks within e-commerce RS. **Value-based Neighbor Selection** (**VNS**) is one technique capable of detecting shilling attacks [10]. This technique can preserve the value dimensions associated with recommendations, which is vital for e-commerce companies. In other words, the recommendations generated by this technique will focus on both profitability and customer relevance. Furthermore, by utilizing users' rating matrix, rating time, and social network analysis output of users' profiles, a Gaussian-Rough neural network can simultaneously leverage both low-order and high-order interactions to detect shilling attacks [57]. This approach advances beyond prior methods that either examine the rating matrix from a single point of view or utilize low-order interactions or high-order interactions.

In addition, an iterative deviation-based user reputation ranking (IDR) method has also been proposed to identify biased or misleading scores in rating items [39]. This method has achieved better results on real sparse datasets with high levels of accuracy and robustness compared to other similar techniques. Using an opinion mining approach to textual reviews that users give for a product allows for identifying push ratings and nuke ratings [15]. In this method, the textual information provided by users is analyzed to determine whether the review is positive, negative, or neutral. Another technique is an ensemble model based on the performance of six machine learning algorithms to detect profile injection attacks [47]. Neighbor Selection with **Variable-Length Partitions** (**VLPNS**) has also been identified to reduce false positive rates by marking suspicious fakers instead of deleting them directly, allowing misclassified normal users to still contribute to the similarity calculation [92]. An unsupervised **multi-modal learning model** (**MMD**), which employs metric learning, has been proposed as another technique to detect professional malicious

users. In another study [101] a stepwise detection method to spot anomalous ratings or attacks, which bypasses the complex problems of similarity calculation and feature extraction, has been developed. Further, a novel detection method, that is, resistant to shilling attacks and profile injection attacks has been developed [100]. According to them, existing features can be summarized into two aspects: rating behavior-based and item distribution-based. Their method formulated the problem as finding a mapping model between rating behavior and item distribution by exploiting the least-squares approximate solution. Based on the train model, they have designed a detector by employing a regressor to detect shilling and profile injection attacks.

*6.1.5 Biased customer feedback.* There are some circumstances when customers provide dishonest feedback. This could be due to the influence of actions of other customers (conformity/social influence bias), the order sequence in which existing online reviews are displayed to a customer (sequential bias), or their tendency to make decisions under the influence of certain emotional states and/or interests [4, 23, 35, 91]. These situations are commonly known as "natural noise", which is different from malicious attacks, which are intentional and aimed at either promoting their products or demoting competitors' products [10]. Biased customer feedback becomes a hazard when it is used as input in the recommendation process, leading to inaccuracies in the product recommendations made to other customers.

Compared to other issues highlighted in this section, the technological solutions to mitigate the impact of this biased feedback within the e-commerce RS context are limited. The InvPref framework [91] is capable of controlling for diverse types of biases to reduce their impact on the recommendation process. Performing sentiment analysis to filter out biased user preferences based on negative polarity scores is also another technique to address biases in feedback [4]. Furthermore, DBT RS (De-biased Tendency) analyzes the bias in product ratings by recalculating the average ratings of products, incorporating user tendencies into the process to counter the issue of biased customer feedback [35].

*6.1.6 Incompetent e-commerce RS.* Some researchers have identified two instances when e-commerce RS do not meeting customer expectations. First is the incompetent RS with functional issues, such as non-consideration of complex user requirements and information overload [45, 87]. Second is the inability of RS to capture changing customer preferences for products over time (e.g., RS fails to capture a customer's changing preferences from smartphones to laptops, leading to irrelevant recommendations). There are several technological solutions to address these issues. First, **Inconsistency Reduction Tools** (**IRTs**) can identify the differences in preference elicitation between customers and other advice sources [45]. Second, the Biased Autoencoder Model (Biased AutoRec) for **collaborative filtering** (**CF**) RS, which is built on the well-known AutoRec CF approach, captures changes in customer preferences [21]. Third, combining **Convolutional Neural Networks** (**CNNs**) with CF to create a hybrid e-commerce RS that leverages behavioral data to improve precision, user interaction, and personalization, ultimately enhancing the prediction of customer interests [48]. In another study [98], a risk-aware recommendation framework has been developed that aims to provide the most suitable recommendations to customers based on their risk attitudes. This framework integrates machine learning and behavioral economics to uncover the risk mechanism behind users' purchasing behaviors. The following section outlines various hazards that can be mitigated through increased customer awareness.

### 6.2 Customer awareness

Customers are an essential stakeholder group in the RS environment. Customers rely on product recommendations from e-commerce retailers to find ideal products that meet their needs. Researchers have highlighted the strong need to increase awareness among customers/the general

public regarding the hazards associated with recommendation technologies, in addition to the technological solutions [13, 72]. These are algorithmic opacity, privacy breaches, and decision biases. Raising customers' awareness of these potential issues, both within and outside the system, can mitigate the adverse effects of the situations highlighted. This section discusses these issues and the solutions in detail.

One of the prominent hazards associated with RS is the users' limited awareness or knowledge of how they receive a particular recommendation. This is referred to as "algorithm opacity," where customers cannot understand how products are recommended to them [13]. The algorithm's opacity could also prevent users from identifying potential biases in the product recommendations presented to them. To overcome this issue, users should be able to understand how and why specific recommendations are provided to them. One technique to address algorithm opacity is the **Conversational Recommender Systems** (hereinafter referred to as **CRS**). CRS have become crucial to address the problem of explanation. CRS are defined as a "software system that supports its users in achieving recommendation-related goals through multi-turn dialogue" [24]. CRS engage users in real time, enabling the generation of recommendations that are closely aligned with their immediate search intentions through an iterative process of dialogue and refinement. The interactive nature of CRS facilitates the collection of granular, user-provided feedback, allowing for a more comprehensive and nuanced understanding of individual preferences beyond observable behavioral data. Furthermore, the integration of explanatory components within the recommendation process enhances users' comprehension of the system's rationale, thereby fostering greater transparency and trust [24]. Collectively, these characteristics contribute to improved recommendations relevance, user satisfaction, and overall system effectiveness.

From a privacy perspective, customer awareness is a crucial aspect in the development of privacy-preserving technologies [25]. Furthermore, this awareness is vital for customers to make an informed decision about the trade-off between the benefits and risks of personalized recommendations [28]. The rationale for this is that when customers are empowered with information, they can regulate their privacy at the desired level. In addition, human decision-making can become biased in certain circumstances. In the e-commerce RS context, there are multiple occasions when these human biases can interfere with customers' decision-making. Two such situations are when customers tend to favor a product simply because it is ranked high (position bias) and when neighboring or related items influence the click-through rate of a target product (neighboring bias) [37]. Enhanced customer awareness is crucial for understanding the impact of these biases, which can subtly influence their decision-making. However, it is noteworthy to mention the dearth of research examining customer awareness or their perceptions of these kinds of biases that cause risks.

Another critical area where customer awareness is often insufficient is the widespread use of A/B testing and online experiments by e-commerce platforms. These experiments are designed to optimize user engagement, recommendation accuracy, or sales conversion rates by exposing different user groups to varying versions of algorithms or interfaces [59]. While such testing has contributed significantly to systems improvement, it raises ethical concerns regarding informed consent, transparency, and user autonomy. Many users are unaware that they are participating in these experiments, which may alter the recommendations they receive and, consequently, influence their purchasing behavior without their explicit knowledge [63, 64]. This lack of awareness may undermine trust in recommendation technologies. Increasing transparency around these experimental practices by informing users when such tests are conducted and offering opt-out options can enhance customer trust and promote a more ethical use of RS [68]. The next section of the review highlights the regulatory developments in several countries aimed at mitigating the effects of hazards associated with e-commerce RS.

### 6.3 Laws and regulations

Personalized recommendations in e-commerce have attracted significant regulatory attention. RS is an essential application of **Artificial Intelligence** (**AI**) that can have implications on fairness, security, privacy, and explainability. Vigorous law enforcement by the government is vital to safeguard customers against hazards associated with e-commerce RS, such as privacy breaches [72]. Accordingly, governments and supranational agencies have introduced laws and regulations to address potential challenges associated with personalized recommendations [20]. This section provides an overview of regulatory measures surrounding personalized recommendations, with a specific focus on the latest developments in Australia, Canada, China, the **European Union** (**EU**), the **United Kingdom** (**UK**), and the **United States** (**US**). These jurisdictions collectively represent a diverse set of regulatory philosophies and major global markets. Together, they capture rights-based frameworks (e.g., the EU's GDPR), market-driven approaches (e.g., the US), and state-led regulatory models (e.g., China), enabling a comprehensive discussion. These countries and regions are also home to leading e-commerce platforms and have been early movers in developing regulations or guidance on algorithmic decision-making and personalization technologies. Their policies not only shape domestic practices but also influence international norms, making them highly relevant for understanding emerging regulatory trends and for identifying best practices that could inform policymaking in other jurisdictions.

*6.3.1 Australia.* The Australian government has shown interest in implementing a governance mechanism to ensure AI is developed and used safely and responsibly in Australia. The Department of Industry, Science, and Resources first published a discussion paper [19][1] seeking public feedback on the governance of AI in Australia. In this discussion paper, AI-enabled personalized recommendations are categorized as "Low Risk." This category is described as having "Minor impacts that are limited, reversible, or brief" [19]. In January 2024, the Australian government published its interim response to this consultation. While the interim response has acknowledged the need for continued work on the regulatory framework for the development and use of AI, it has also emphasized the importance of not hindering the use of low-risk AI. In December 2022, the Safety Commissioner (an independent statutory office responsible for implementing and enforcing the Online Safety Act, 2021) released a Position Statement,[2] highlighting the potential for RS to enhance users' experiences while also posing risks by amplifying harmful content. The statement advocates for transparency, user control, and safety by design, aiming to balance the benefits of personalized content with the need to mitigate potential harm.

*6.3.2 Canada.* In Canada, the Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems [11][3] seek voluntary commitment from developers and managers to achieve several key outcomes. The outcomes are Accountability, Safety, Fairness and Equity, Transparency, Human Oversight and Monitoring, and Validity and Robustness. The code does not categorize different AI-based applications; hence, there are no specific guidelines applicable to personalized product recommendations. However, the code outlines measures that are broadly applicable to a range of AI systems, which system developers and managers should fulfill voluntarily. As of now, there are 31 signatories to the code, indicating widespread acceptance. This code aims to provide a bridge to the AI and Data Act, which has yet to be recognized as a statute in Canada.

[1] https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public_assets/Safe-and-responsible-AI-in-Australia-discussion-paper.pdf
[2] https://www.esafety.gov.au/industry/tech-trends-and-challenges/recommender-systems-and-algorithms
[3] https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act

Table 8. Provisions and their Implications for Hazard Mitigation - China

| Articles and clauses | Summary of the provision | Implications for hazard mitigation in e-commerce RS |
|---|---|---|
| Article 6 | Providers must align their algorithms with mainstream values, actively disseminate positive content, and prevent the spread of information prohibited by laws and regulations. | This provision can help restrict the spread of misinformation, which may include inappropriate consumption practices that harm the well-being of customers and society. |
| Article 7 | Providers are required to establish management systems and technical measures for algorithmic security, including user registration, information verification, data security, and protection of personal information. | This provision encourages companies to be vigilant in protecting customer data. |
| Article 8 | Providers must regularly review and assess their algorithms to ensure they do not induce user addiction or excessive spending. | This provision aims to ensure customer well-being in relation to their consumption habits. |
| Article 16 | Providers must transparently and understandably inform users about the basic principles, purposes, and main operating mechanisms of their algorithmic recommendation services. | This provision guides providers to educate customers, which could help reduce algorithmic opacity. |
| Article 17 | Users should be given options to disable algorithmic recommendations and manage user tags that influence personalized content. If a user opts to disable such services, the provider must cease providing them immediately. | This provision empowers customers to decide whether they require personalized recommendations. |

*6.3.3 China.* In China, extensive guidelines for regulating personalized recommendations are provided in the Provisions on the Management of Algorithmic Recommendations in Internet Information Services. These provisions are drafted in accordance with the Cybersecurity, Data Security, and Personal Information Protection Laws of China.[4] Table 8 outlines key provisions from China's regulatory framework that target hazard mitigation in e-commerce RS. These regulations emphasize algorithmic alignment with societal values (Article 6), aiming to curb the spread of harmful or misleading content, including inappropriate consumption practices. A strong focus is placed on algorithmic security and data protection (Article 7), requiring providers to implement robust safeguards for user data. This is complemented by Article 8, which mandates regular reviews to prevent excessive user engagement or spending, thereby protecting customer well-being and promoting healthier consumption behavior. In terms of transparency and user awareness, Article 16 requires platforms to communicate how their recommendation algorithms operate. This helps reduce algorithmic opacity and builds user trust. Additionally, Article 17 enhances user autonomy by allowing individuals to opt out of personalized recommendations and manage their profiling data. Together, these provisions reflect a regulatory approach centered on social responsibility, user protection, and transparency. They encourage e-commerce platforms to design RS that are both secure and aligned with ethical consumption standards.

*6.3.4 European Union.* Several legislations and regulations cover the personalized recommendations domain. Out of these, the **Digital Service Act (DSA)**[5] is the most prominent legislation, which was introduced in 2022 and came into full effect in 2024. The DSA applies to all digital services that connect customers to goods, services, and content.

[4] https://www.chinalawtranslate.com/en/algorithms/
[5] European Commission, Digital Services Act

Table 9. Provisions and their Implications for Hazard Mitigation - EU

| Articles and clauses | The provisions | Implications for hazard mitigation in e-commerce RS |
|---|---|---|
| Article 27 and Clause 70 | VLOPs must inform users about how RS impact the display and prioritization of information. VLOPs should clearly explain the criteria and identify the factors RS uses to display information in an easily understandable way. | This provision can mitigate the impact of algorithmic opacity, enabling customers to evaluate the value of the recommendations they receive. |
| Article 34 and Clause 84 | VLOPs should assess systemic risks by examining systems, including RS, along with their data collection and usage practices. They should focus on system risks such as the spread of misleading or deceptive information, including disinformation. | These provisions can help restrict the spread of misinformation, which may include inappropriate consumption practices that harm the well-being of customers and society. |
| Article 35 and Clause 88 | VLOPs should regularly test and, if needed, adjust their RS (i.e., correcting the criteria used in RS) to mitigate adverse effects, such as harmful personalized recommendations. They are also required to cooperate with other platforms through codes of conduct or self-regulatory measures, particularly when dealing with shared risks, such as disinformation campaigns. | These provisions direct companies to continuously assess the possibility of RS causing risks to customers. Cooperation among platforms will be a deterrent to the spreading of misinformation, especially in the case of social media-based product recommendations. |
| Article 38 and Clause 94 | VLOPs and search engines must assess and adjust their RS to prevent or minimize biases, especially when those biases could lead to discrimination against vulnerable individuals. Such adjustments must comply with data protection laws, particularly when dealing with special categories of personal data. Furthermore, providers should offer users easy access options for recommendations that do not rely on profiling, in accordance with data protection regulations. | This provision addresses two key issues. Minimizing biases in recommendations will ensure fair treatment for all types of customers in the e-commerce context. Enabling non-profiled recommendations will provide users with more autonomy in their decision-making and enhanced security for their personal information. |
| Article 40 and Clause 96 | The Digital Services Coordinator or the Commission may request access to specific data, including algorithm-related data, to monitor and assess compliance with the regulation by VLOPs. This can include data needed to determine risks and potential harm, as well as information on the accuracy, functioning, and testing of algorithms in content moderation, RS, and advertising systems. | This provision empowers the regulators to assess whether the RS are functioning in compliance with the regulations, especially in a manner that does not cause harm to customers. |
| Article 44 and Clause 102 | The regulation encourages the development of voluntary industry standards for various technical procedures, including RS. | Industry-specific standards will provide an opportunity for companies to consider the specific hazards they face in the e-commerce RS context. For example, balancing profitability and the diversity of product recommendations simultaneously. |

Table 9 highlights key EU regulatory provisions aimed at mitigating risks in e-commerce RS, particularly on **Very Large Online Platforms** (**VLOPs**). A central focus is on transparency, requiring platforms to clearly explain how RS influence content visibility (Article 27), helping users better understand and evaluate recommendations. The regulations also address systemic risks, such as the spread of misleading or harmful content (Article 34), and require ongoing evaluation and adjustments to RS to reduce adverse effects (Article 35). This proactive approach encourages platforms to collaborate and maintain accountability.

Bias reduction and user autonomy are further emphasized (Article 38), promoting fairer treatment and offering alternatives to profiling-based recommendations. Regulatory oversight is strengthened by granting access to algorithmic data (Article 40), enabling enforcement bodies to assess compliance and prevent harm. Finally, the push for voluntary industry standards (Article 44) supports the development of context-specific best practices, enabling companies to balance their business goals with ethical and inclusive design recommendations. Collectively, these measures create a more transparent, fair, and user-aligned RS environment in the EU.

*6.3.5 United Kingdom.* In the UK, a white paper titled "A pro-innovation approach to AI regulation" [86] was published in 2023 to guide the regulation of AI. In this article, the UK government has clearly stated that AI offers both challenges and opportunities. To maximize the benefits of AI applications, this white paper recommends adopting an evidence-based approach to learning from experience and continuously adapting to develop the most effective regulatory regime. In this white paper, content recommendation is identified as an application of AI that carries both risks and enormous benefits in our day-to-day lives.

Although not explicitly referred to as RS, several other pieces of legislation can impact the way RS are used in the UK. The Online Safety Act (2023)[6] imposes new requirements on digital platforms, particularly when they impact user safety or involve content that could be harmful. The Digital Markets, Competition, and Consumer Act (2024)[7] aims to curtail misleading practices by digital platforms with significant market power. For example, the act includes several misleading actions, such as providing false or misleading information and presenting it deceptively. These existing legislations, along with the proposed AI regulation framework, will strengthen the regulatory framework for the use of AI-based applications, such as e-commerce RS.

*6.3.6 United States.* In the US, there are no laws specifically regulating RS. However, several federal laws have implications for the implementation of RS. The Algorithmic Accountability Act (2023)[8] requires companies to conduct assessments evaluating the impacts of their automated decision systems, augmented critical decision processes, or the impact assessments of such systems or processes in areas such as accuracy, reliability, fairness (including bias and non-discrimination), explainability, and other relevant factors. In addition, Section 230[9] of the Communications Decency Act (1996) provides immunity to online platforms (including those involved in commerce) from liability for third-party or user-generated content. Section 230 provides immunity to online platforms both for third-party content on their services and for the removal of specific categories of content. This was meant to nurture emerging internet businesses while also incentivizing them to regulate harmful online content. However, over time, there has been a divergence from this original purpose. This is evidenced by reduced incentives for online platforms to address illicit activities on their services, which has left them free to moderate lawful content without transparency and accountability. While many have called for the repeal of this section, citing the growing online harms, Section 230 is in effect.

In summary, when analyzing the laws, they can be placed in a continuum based on their level of applicability to personalized recommendations, with some statutes directly impacting them, while others do not. For instance, while Canada has minimal regulatory coverage, the EU and China have extensive regulations. However, the Canadian Code of Conduct recognizes the roles played by developers and managers in the execution of AI-enabled systems, recognizing RS as multi-stakeholder systems. It is encouraging to see the depth of the DSA, which has gone into specific

[6] https://www.legislation.gov.uk/ukpga/2023/50
[7] https://www.legislation.gov.uk/ukpga/2024/13/enacted
[8] https://www.congress.gov/bill/118th-congress/senate-bill/2892/text
[9] https://www.congress.gov/crs-product/R46751

Table 10. Laws and their Applicability to Specific Hazards

| Country/Region | Laws/Regulations/Guidelines | Specific hazards |
|---|---|---|
| Australia | Discussion paper on the governance of AI | No hazards addressed directly. |
| | Position Statement by the eSafety Commissioner | No hazards addressed directly. |
| Canada | Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems | No hazards addressed directly. |
| China | Provisions on the Management of Algorithmic Recommendations in Internet Information Service | – Spreading of misinformation. |
| | | – Data protection. |
| | | – Reducing algorithmic opacity. |
| | | – Minimizing biases in recommendations. |
| EU | DSA | – Algorithmic opacity. |
| | | – Spreading of misinformation. |
| | | – Minimizing biases in recommendations. |
| | | – Enhanced user control over personal information. |
| UK | The Online Safety Act (2023) | – Spreading of misinformation. |
| | | – Data protection. |
| | The Digital Markets, Competition, and Consumer Act (2024) | – Data protection. |
| US | The Algorithmic Accountability Act (2023) | – Spreading of misinformation. |
| | | – Data protection. |
| | | – Minimizing biases in recommendations. |

details of different areas of RS technology applications that could cause risks to the users. The EU regulations will provide an essential impetus for other countries to have more robust and broader regulatory frameworks to govern the commercial applications of RS and related technologies. Table 10 provides an overview of the applicability of the laws under consideration to specific hazards discussed in this review.

## 7 Discussion

This review aimed to address the research question of what solutions have been proposed to mitigate the hazards associated with e-commerce RS, and how these can be categorized into technological, customer awareness, and regulatory dimensions to form a more effective response. Accordingly, we have identified a range of solutions aimed at mitigating the hazards associated with e-commerce RS and categorized them as mentioned above. Our analysis revealed that these solutions have progressed to varying extents. Notably, technological solutions addressing different hazards have received unequal levels of attention from researchers. A substantial proportion of research efforts have been dedicated to mitigating issues such as algorithmic biases in product recommendations, privacy breaches, fake reviews, and adversarial attacks. However, comparatively less research has focused on addressing inefficiencies in e-commerce RS and mitigating the impact of human decision-making biases.

This difference in research may be due to the frequency and ease with which different hazards can be identified. For example, biases in product recommendations are often more transparent and easier to measure, which makes them more suitable for systematic analysis and technological solutions. On the other hand, biases in human decision-making tend to be more subjective and

context-dependent, making them harder to identify and rectify precisely. Therefore, future research should focus more on these less-studied areas to develop a more comprehensive and balanced approach to improving e-commerce risk management.

When discussing different techniques developed to counter popularity bias, it was evident that they would have a multitude of outcomes. Providing prominence to less popular products (i.e., long-tail items) may also lead to enhanced diversity in product recommendations and fairness to less popular producers. In addition, it will have positive social implications by mitigating issues such as filter bubbles that can force customers into extreme consumption habits [7]. Although different types of biases in product recommendations are discussed in the research, it is questionable to what extent customers are aware of those biases. However, when examining regulatory measures (i.e., EU regulations) that are in place to mitigate biases in recommendations, it becomes clear that government agencies are aware of the adverse effects of biased recommendations. The cold-start problem is discussed in the RS literature as a key hazard concerning less-known or first-time users. In the e-commerce context, this refers to customers who are new to a platform. It would be interesting to investigate the real-life experiences of such customers to understand the impact of the cold-start problem on them. Again, when examining regulations, especially in the case of the EU, provisions on fair treatment for all types of customers should provide some assurance in curtailing such experiences.

Privacy breaches have attracted significant attention among e-commerce RS researchers. The privacy-personalization paradox is a key topic in the e-commerce RS context, as the increasing reliance on customer data to enhance recommendations can raise privacy concerns if not managed properly [73]. Unfortunately, privacy risks in e-commerce RS have led customers to engage in sub-optimal behaviors such as withholding personal information [77], providing false data [43, 67], and engaging in free-riding behavior where they browse for products online but ultimately make their purchases offline [99]. Such behaviors pose challenges for e-commerce RS because both implicit and explicit customer feedback, essential for improving recommendations, can only be effectively gathered when more customers actively engage in online shopping rather than merely searching for information. Enhancing customer awareness to complement evolving regulatory requirements is essential for e-commerce companies to foster informed customer engagement in the personalization space. As previously discussed, informed customers will be better equipped to make decisions regarding the tradeoff between privacy and benefits.

Fake reviews and ratings have garnered significant attention from researchers, as evidenced by the wide range of technological solutions proposed to mitigate their adverse effects. From a customer's perspective, identifying these deceptive practices remains highly challenging. Further research on fake reviews and ratings is essential for fostering fairness in e-commerce RS for both customers and companies. Ensuring the authenticity of reviews and ratings benefits companies whose products genuinely meet customer expectations. Additionally, reliable reviews contribute to more effective product recommendations, ultimately enhancing the overall customer experience.

Moreover, research has indicated that various unidentified biases exist within the feedback loop of RS. Advancements in general debiasing techniques have the potential to address these varying types of biases [91]. However, this remains a complex challenge, as human biases are inherently subjective and often difficult to detect. Raising customer awareness about these biases could help them become more critical and discerning when evaluating recommendations.

Although CRS are primarily a technological solution, they have also been discussed as a means of increasing customer awareness due to their explanatory capabilities. CRS can engage customers in a dialogue to understand their preferences better and deliver more effective, personalized product recommendations. This represents a significant technological advancement, empowering customers with greater knowledge about why specific products are recommended to them. However, one major challenge associated with CRS involves privacy and ethical concerns. CRS collect extensive

data through customer interactions, which may introduce challenges in real-world implementation. Additionally, the explanations provided by CRS could compromise customer privacy, particularly if personal information is exposed to unintended audiences, such as individuals observing the user's screen [28].

Most of the countries cited in this article have either introduced or are in the process of introducing regulations to govern personalized recommendations. As these regulations are still in their early stages, it is premature to assess their implementation and effectiveness. However, it will be essential to observe how the e-commerce companies navigate these heightened regulatory requirements. Given that most of the world's leading e-commerce companies operate as multinational corporations, the applicability and enforcement of these across jurisdictions warrant close examination.

## 8 Implications and Future Work

This review emphasizes the need for a more proactive and integrated research agenda in the domain of e-commerce RS. While technological, regulatory, and awareness-driven solutions offer distinct strengths, future work should prioritize how these domains interact and co-evolve to ensure systemic resilience. Rather than treating them in isolation, future research should explore frameworks that enable the effective coordination of these approaches, particularly in the context of increasingly complex and AI-driven RS.

One promising direction is the ongoing development of CRS. Beyond their current capabilities, there is scope to explore their potential in delivering multi-modal, context-aware interactions that integrate voice, text, and visual cues. Advancing these systems will involve both technical innovation and rigorous user testing to ensure that improvements in adaptability do not compromise performance or user privacy. Research should also examine how CRS can provide tailored explanations without inadvertently exposing sensitive information, which is a challenge that grows as CRS become more pervasive.

A significant research gap exists in the lack of empirical studies on customer perceptions of RS-related hazards. Future research should employ structured methods, such as the psychometric paradigm, to systematically capture how users understand and respond to risks, including algorithmic bias, privacy loss, and fake reviews. These insights will help improve risk communication strategies and educational tools, enabling customers to engage more thoughtfully with personalized recommendations.

From a regulatory perspective, the changing and inconsistent global policy environment presents both challenges and opportunities. Comparative studies are needed to assess how businesses in jurisdictions with differing regulatory intensities, such as the EU and Australia, adapt their RS practices. More broadly, future research should capture industry perspective, especially in countries with emerging or absent regulatory frameworks, to ensure that future policies are practical, enforceable, and responsive to industry realities.

Notably, the EU DSA explicitly encourages academic engagement, particularly with VLOPs. Clause 96 of the DSA calls for research into how systemic risks develop and grow, emphasizing the need for continuous monitoring for hazards as RS technologies develop. This directive implies a broader research imperative: to design empirical, longitudinal studies that track emerging risks in real time and assess the adequacy of current safeguards.

In summary, future work should focus on:

(1) Enhancing the adaptability and privacy-preserving capabilities of CRS.
(2) Empirically investigating customer perceptions of RS hazards.
(3) Integrating business perspectives into regulatory development.
(4) Conducting longitudinal risk assessments in response to regulatory calls, especially in AI-enabled RS environments.

Addressing these areas will be essential for building RS that are not only effective but also equitable, transparent, and resilient to both current and emerging risks.

## 9 Conclusion

In this review, we explored solutions to mitigate the hazards associated with e-commerce product recommendations. We categorized them into three categories, namely technological solutions, increasing customer awareness, and implementing laws and regulations. We provided examples within each of these categories to illustrate the current state in the e-commerce RS domain. Our discussion highlighted that these three solutions must function in tandem to address the risks posed by e-commerce RS effectively. Additionally, we outlined several future research directions that could provide valuable insights for researchers in this field. In summary, while significant progress has been made across these three domains, ongoing research and regulatory efforts are still crucial, especially given the rapid advancements in AI-driven recommendation technologies.

## References


[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Pizzato. 2020. Multistakeholder recommendation: Survey and research directions. *User Modeling and User-Adapted Interaction* 30, 1 (2020), 127–158.

[2] Gediminas Adomavicius, Jesse C. Bockstedt, Shawn P. Curley, and Jingjing Zhang. 2013. Do recommender systems manipulate consumer preferences? A study of anchoring effects. *Information Systems Research* 24, 4 (2013), 956–975. DOI: https://doi.org/10.1287/isre.2013.0497

[3] Ghazaleh Aghili, Mehdi Shajari, Shahram Khadivi, and Mohammad Amin Morid. 2011. Using genre interest of users to detect profile injection attacks in movie recommender systems. In *Proceedings of the 2011 10th International Conference on Machine Learning and Applications and Workshops*, Vol. 1. IEEE, 49–52. DOI: https://doi.org/10.1109/ICMLA.2011.151

[4] Adeel Ahmed, Khalid Saleem, Osman Khalid, Jiechao Gao, and Umer Rashid. 2022. Trust-aware denoising autoencoder with spatial-temporal activity for cross-domain personalized recommendations. *Neurocomputing* 511 (2022), 477–494. DOI: https://doi.org/10.1016/j.neucom.2022.09.023

[5] Pegah Malekpour Alamdari, Nima Jafari Navimipour, Mehdi Hosseinzadeh, Ali Asghar Safaei, and Aso Darwesh. 2020. A systematic study on the recommender systems in the E-commerce. *Ieee Access* 8 (2020), 115694–115716. DOI: https://doi.org/10.1109/ACCESS.2020.3002803

[6] S. Annamalai, N. Sangeetha, M. Kumaresan, Dommaraju Tejavarma, Gandhodi Harsha Vardhan, and A. Suresh Kumar. 2025. Application domains of federated learning. *Model Optimization Methods for Efficient and Edge AI: Federated Learning Architectures, Frameworks and Applications* (2025), 127–144. Retrieved from https://onlinelibrary.wiley.com/doi/chapter-epub/10.1002/9781394219230.ch7

[7] Ricardo Baeza-Yates and Giovanni Delnevo. 2022. Exploration trade-offs in web recommender systems. In *Proceedings of the 2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 1–9. DOI: https://doi.org/10.1109/BigData55660.2022.10325847

[8] Zeynep Batmaz, Ali Yurekli, Alper Bilge, and Cihan Kaleli. 2019. A review on deep learning for recommender systems: Challenges and remedies. *Artificial Intelligence Review* 52, 1 (2019), 1–37. DOI: https://doi.org/10.1007/s10462-018-9654-y

[9] Alon Ben Horin and Tamir Tassa. 2021. Privacy preserving collaborative filtering by distributed mediation. In *Proceedings of the 15th ACM Conference on Recommender Systems*. 332–341. DOI: https://doi.org/10.1145/3460231.3474251

[10] Yuanfeng Cai and Dan Zhu. 2019. Trustworthy and profit: A new value-based neighbor selection method in recommender systems under shilling attacks. *Decision Support Systems* 124 (2019), 113112. DOI: https://doi.org/10.1016/j.dss.2019.113112

[11] Canadian Government. 2024. Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems, 2024. Retrieved August 30, 2024 from https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems

[12] Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. 2023. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems* 41, 3 (2023), 1–39.

[13] Yuyao Chen. 2022. Analysis on the impact of recommender systems on consumer decision: Making on China's online shopping platforms. In *Proceedings of the 6th International Conference on E-Commerce, E-Business and E-Government*. 29–33. DOI: https://doi.org/10.1145/3537693.3537734

[14] Akanksha Bansal Chopra and Veer Sain Dixit. 2020. Balanced accuracy of collaborative recommender system. In *Proceedings of the ICT Systems and Sustainability: Proceedings of ICT4SD 2020, Volume 1*. Springer, 341–356. DOI: https://doi.org/10.1007/978-981-15-8289-9_32

[15] Akanksha Bansal Chopra and Veer Sain Dixit. 2023. Detecting biased user-product ratings for online products using opinion mining. *Journal of Intelligent Systems* 32, 1 (2023), 20229030. DOI: https://doi.org/10.1515/jisys-2022-9030

[16] Chen-Yao Chung, Ping-Yu Hsu, and Shih-Hsiang Huang. 2013. $\beta$P: A novel approach to filter out malicious rating profiles from recommender systems. *Decision Support Systems* 55, 1 (2013), 314–325. DOI: https://doi.org/10.1016/j.dss.2013.01.020

[17] Shay David and Trevor John Pinch. 2005. Six degrees of reputation: The use and abuse of online review and recommendation systems. *Available at SSRN 857505* (2005). DOI: https://doi.org/10://doi.org/10.5210/fm.v11i3.1315

[18] Wei Deng, Yong Shi, Zhengxin Chen, Wikil Kwak, and Huimin Tang. 2020. Recommender system for marketing optimization. *World Wide Web* 23, 3 (2020), 1497–1517. DOI: https://doi.org/10.1007/s11280-019-00738-1

[19] Department of Industry,Science and Resource, Australia. 2023. Safe and responsible AI in Australia. Retrieved August 8, 2023 from https://consult.industry.gov.au/supporting-responsible-ai

[20] Tommaso Di Noia, Nava Tintarev, Panagiota Fatourou, and Markus Schedl. 2022. Recommender systems under European AI regulations. *Communications of the ACM* 65, 4 (2022), 69–73. DOI: https://doi.org/10.1145/3512728

[21] Runliang Dou, Oguzhan Arslan, and Ce Zhang. 2021. Biased autoencoder for collaborative filtering with temporal signals. *Expert Systems with Applications* 186 (2021), 115775. DOI: https://doi.org/10.1016/j.eswa.2021.115775

[22] Zekeriya Erkin, Michael Beye, Thijs Veugen, and Reginald L. Lagendijk. 2012. Privacy-preserving content-based recommender system. In *Proceedings of the on Multimedia and security*. 77–84. DOI: https://doi.org/10.1145/2361407.2361420

[23] Enes Eryarsoy and Selwyn Piramuthu. 2014. Experimental evaluation of sequential bias in online customer reviews. *Information & Management* 51, 8 (2014), 964–971. DOI: https://doi.org/10.1016/j.im.2014.09.001

[24] Matteo Fabbri. 2023. Social influence for societal interest: a pro-ethical framework for improving human decision making through multi-stakeholder recommender systems. *Ai & Society* 38, 2 (2023), 995–1002. DOI: https://doi.org/10.1007/s00146-022-01467-2

[25] Eva-Patricia Fernández Manzano and María Isabel González Vasco. 2018. Analytic surveillance: Big data business models in the time of privacy awareness. *El Profesional de la Información* 27, 2 (2018), 402–409. DOI: https://doi.org/10.3145/epi.2018.mar.19

[26] Daniel Fleder and Kartik Hosanagar. 2009. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science* 55, 5 (2009), 697–712. DOI: https://doi.org/10.1287/mnsc.1080.0974

[27] Remo Frey, Dominic Wörner, and Alexander Ilic. 2016. Collaborative filtering on the blockchain: A secure recommender system for e-commerce. (2016). Retrieved from https://aisel.aisnet.org/amcis2016/ISSec/Presentations/36

[28] Arik Friedman, Bart P. Knijnenburg, Kris Vanhecke, Luc Martens, and Shlomo Berkovsky. 2015. Privacy aspects of recommender systems. In *Proceedings of the Recommender Systems Handbook*. Springer, 649–688. DOI: https://doi.org/10.1007/978-1-4899-7637-6_19

[29] Isaac J. Gabriel and Easwar Nyshadham. 2008. A cognitive map of people's online risk perceptions and attitudes: An empirical study. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. IEEE, 274–274. DOI: https://doi.org/10.1109/HICSS.2008.6

[30] Chen Gao, Yu Zheng, Nian Li, Yinfeng Li, Yingrong Qin, Jinghua Piao, Yuhan Quan, Jianxin Chang, Depeng Jin, Xiangnan He, et al. 2023. A survey of graph neural networks for recommender systems: Challenges, methods, and directions. *ACM Transactions on Recommender Systems* 1, 1 (2023), 1–51. DOI: https://doi.org/10.1145/3568022

[31] Yingqiang Ge, Shuchang Liu, Zuohui Fu, Juntao Tan, Zelong Li, Shuyuan Xu, Yunqi Li, Yikun Xian, and Yongfeng Zhang. 2024. A survey on trustworthy recommender systems. *ACM Transactions on Recommender Systems* 3, 2 (2024), 1–68.

[32] Yingqiang Ge, Shuyuan Xu, Shuchang Liu, Zuohui Fu, Fei Sun, and Yongfeng Zhang. 2020. Learning personalized risk preferences for recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 409–418. DOI: https://doi.org/10.1145/3397271.3401056

[33] Steven Glover and Izak Benbasat. 2010. A comprehensive model of perceived risk of e-commerce transactions. *International Journal of Electronic Commerce* 15, 2 (2010), 47–78. DOI: https://doi.org/10.2753/JEC1086-4415150202

[34] Dhananjaya Gm, R. H. Goudar, Anjanabhargavi A. Kulkarni, Vijayalaxmi N. Rathod, and Geetabai S. Hukkeri. 2024. A digital recommendation system for personalized learning to enhance online education: A review. *IEEE Access* 12 (2024), 34019–34041.

[35] Mukkamula Venu Gopalachari. 2018. DBT recommender: Improved trustworthiness of ratings through de-biasing tendency of users. *International Journal of Intelligent Engineering and Systems* 11, 2 (2018), 85–92. DOI: https://doi.org/10.22266/ijies2018.0430.10

[36] Camille Grange, Izak Benbasat, and Andrew Burton-Jones. 2019. With a little help from my friends: Cultivating serendipity in online shopping environments. *Information & Management* 56, 2 (2019), 225–235.

[37] Yulong Gu, Zhuoye Ding, Shuaiqiang Wang, Lixin Zou, Yiding Liu, and Dawei Yin. 2020. Deep multifaceted transformers for multi-objective ranking in large-scale e-commerce recommender systems. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2493–2500. DOI: https://doi.org/10.1145/3340531.3412697

[38] Liang Hu, Longbing Cao, Jian Cao, Zhiping Gu, Guandong Xu, and Jie Wang. 2017. Improving the quality of recommendations for users and items in the tail of distribution. *ACM Transactions on Information Systems (TOIS)* 35, 3 (2017), 1–37. DOI: https://doi.org/10.1145/3052769

[39] Jia-Tao Huang, Hong-Liang Sun, Xiao-Fei Chen, Xiao-lin Liu, and Jie Cao. 2021. An iterative deviation-based ranking method to evaluate user reputation in online rating Systems. In *Proceedings of the 2021 4th international conference on data science and information technology*. 15–21. DOI: https://doi.org/10.1145/3478905.3478909

[40] Dietmar Jannach and Christine Bauer. 2020. Escaping the McNamara fallacy: Towards more impactful recommender systems research. *Ai Magazine* 41, 4 (2020), 79–95. DOI: https://doi.org/10.1609/aimag.v41i4.5312

[41] Dietmar Jannach, Pearl Pu, Francesco Ricci, and Markus Zanker. 2022. Recommender systems: Trends and frontiers. *Ai Magazine* 43, 2 (2022), 145–150. DOI: https://doi.org/10.1002/aaai.12050

[42] Neera Jeyamohan, Xiaomin Chen, and Nauman Aslam. 2019. Local differentially private matrix factorization for recommendations. In *Proceedings of the 2019 13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*. IEEE, 1–5. DOI: https://doi.org/10.1109/SKIMA47702.2019.8982536

[43] Somayeh Moghaddam, Zadeh Kashani, and Javad Hamidzadeh. 2020. Feature selection by using privacy-preserving of recommendation systems based on collaborative filtering and mutual trust in social networks. *Soft Computing* 24, 15 (2020), 11425–11440. DOI: https://doi.org/10.1007/s00500-019-04605-z

[44] Eranjana Kathriarachchi, Shafiq Alam, Kasuni Weerasinghe, and David Pauleen. 2024. Risks of e-commerce recommender systems: A scoping review. *Australasian Journal of Information Systems* 28 (2024), 1–42. DOI: https://doi.org/10.3127/ajis.v28.4869

[45] Hongki Kim, Izak Benbasat, and Hasan Cavusoglu. 2017. Online consumers' attribution of inconsistency between advice sources. (2017). Retrieved from https://www.proceedings.com/37640.html

[46] Dede Kiswanto, Dade Nurjanah, and Rita Rismala. 2018. Fairness aware regularization on a learning-to-rank recommender system for controlling popularity bias in e-commerce domain. In *Proceedings of the 2018 International Conference on Information Technology Systems and Innovation (ICITSI)*. IEEE, 16–21. DOI: https://doi.org/10.1109/ICITSI.2018.8696023

[47] Ashish Kumar, Deepak Garg, and Prashant Singh Rana. 2015. Ensemble approach to detect profile injection attack in recommender system. In *Proceedings of the 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 1734–1740. DOI: https://doi.org/10.1109/ICACCI.2015.7275864

[48] S. Vinodh Kumar, P Kumar, and S. Saravanan. 2025. Enhancing E-commerce using fashion recommendation system. In *Proceedings of the 2025 International Conference on Visual Analytics and Data Visualization (ICVADV)*. IEEE, 864–868. DOI: https://doi.org/10.1109/ICVADV63329.2025.10961349

[49] Yen-Hsien Lee, Paul Jen-Hwa Hu, Tsang-Hsiang Cheng, and Ya-Fang Hsieh. 2012. A cost-sensitive technique for positive-example learning supporting content-based product recommendations in B-to-C e-commerce. *Decision Support Systems* 53, 1 (2012), 245–256. DOI: https://doi.org/10.1016/j.dss.2012.01.018

[50] Chengzhe Li. 2023. Cold start recommendation based on cross-domain sharing knowledge transfer learning. In *Proceedings of the 2023 International Conference on Power, Electrical Engineering, Electronics and Control (PEEEC)*. IEEE, 911–915. DOI: https://doi.org/10.1109/PEEEC60561.2023.00178

[51] Gesu Li, Guisheng Yin, Jishen Yang, and Fukun Chen. 2021. SDRM-LDP: A recommendation model based on local differential privacy. *Wireless Communications and Mobile Computing* 2021, 1 (2021), 6640667. DOI: https://doi.org/10.1155/2021/6640667

[52] Seth Siyuan Li and Elena Karahanna. 2015. Online recommendation systems in a B2C E-commerce context: A review and future directions. *Journal of the Association for Information Systems* 16, 2 (2015), 2.

[53] Sitikantha Mallik and Abhaya Kumar Sahoo. 2019. A comparison study of different privacy preserving techniques in collaborative filtering based recommender system. In *Computational Intelligence in Data Mining: Proceedings of the International Conference on ICCIDM 2018*. Springer, 193–203. DOI: https://doi.org/10.1007/978-981-13-8676-3_17

[54] Daniel Mican, Dan-Andrei Sitar-Tăut, and Ovidiu-Ioan Moisescu. 2020. Perceived usefulness: A silver bullet to assure user data availability for online recommendation systems. *Decision Support Systems* 139 (2020), 113420. DOI: https://doi.org/10.1016/j.dss.2020.113420

[55] Silvia Milano, Mariarosaria Taddeo, and Luciano Floridi. 2020. Recommender systems and their ethical challenges. *Ai & Society* 35, 4 (2020), 957–967. DOI: https://doi.org/10.1007/s00146-020-00950-y

[56] Hanyi Min, Baojiang Yang, David G. Allen, Alicia A. Grandey, and Mengqiao Liu. 2024. Wisdom from the crowd: Can recommender systems predict employee turnover and its destinations? *Personnel Psychology* 77, 2 (2024), 475–496.

[57] R. Moradi and H. Hamidi. 2023. A new mechanism for detecting shilling attacks in recommender systems based on social network analysis and Gaussian rough neural network with emotional learning. *International Journal of Engineering* 36, 2 (2023), 321–334. DOI: https://doi.org/10.5829/IJE.2023.36.02B.12

[58] Michael Sean Murphy. 2011. Notes toward a politics of personalization. In *Proceedings of the 2011 iConference*. 546–551. DOI: https://doi.org/10.1145/1940761.1940836

[59] Preetam Nandy, Divya Venugopalan, Chun Lo, and Shaunak Chatterjee. 2021. A/b testing for recommender systems in a two-sided marketplace. *Advances in Neural Information Processing Systems* 34 (2021), 6466–6477. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2021/hash/32e19424b63cc63077a4031b87fb1010-Abstract.html

[60] Sabina-Cristiana Necula and Vasile-Daniel Păvăloaia. 2023. AI-driven recommendations: A systematic review of the state of the art in e-commerce. *Applied Sciences* 13, 9 (2023), 5531. DOI: https://doi.org/10.3390/app13095531

[61] Ke Niu, Xiangyu Zhao, Fangfang Li, Ning Li, Xueping Peng, and Wei Chen. 2019. UTSP: User-based two-step recommendation with popularity normalization towards diversity and novelty. *IEEE Access* 7 (2019), 145426–145434. DOI: https://doi.org/10.1109/ACCESS.2019.2939945

[62] Umberto Panniello, Shawndra Hill, and Michele Gorgoglione. 2016. The impact of profit incentives on the relevance of online recommendations. *Electronic Commerce Research and Applications* 20 (2016), 87–104. DOI: https://doi.org/10.1016/j.elerap.2016.10.003

[63] Dimitris Paraschakis. 2016. Recommender systems from an industrial and ethical perspective. In *Proceedings of the 10th ACM conference on recommender systems*. 463–466. DOI: https://doi.org/10.1145/2959100.2959101

[64] Dimitris Paraschakis. 2017. Towards an ethical recommendation framework. In *Proceedings of the 2017 11th International Conference on Research Challenges in Information Science (RCIS)*. IEEE, 211–220. DOI: https://doi.org/10.1109/RCIS.2017.7956539

[65] Yanni Ping, Yang Li, and Jiaxin Zhu. 2025. Beyond accuracy measures: The effect of diversity, novelty and serendipity in recommender systems on user engagement. *Electronic Commerce Research* 25, 3 (2025), 2177–2204. DOI: https://doi.org/10.1007/s10660-024-09813-w

[66] Luiz Pizzato, Tomasz Rej, Joshua Akehurst, Irena Koprinska, Kalina Yacef, and Judy Kay. 2013. Recommending people to people: The nature of reciprocal recommenders with a case study in online dating. *User Modeling and User-Adapted Interaction* 23, 5 (2013), 447–488.

[67] Huseyin Polat and Wenliang Du. 2005. Privacy-preserving collaborative filtering. *International Journal of Electronic Commerce* 9, 4 (2005), 9–35. DOI: https://doi.org/10.1080/10864415.2003.11044341

[68] Andrea Polonioli, Riccardo Ghioni, Ciro Greco, Prathm Juneja, Jacopo Tagliabue, David Watson, and Luciano Floridi. 2023. The ethics of online controlled experiments (A/B Testing). *Minds and Machines* 33, 4 (2023), 667–693. DOI: https://doi.org/10.1007/s11023-023-09644-y

[69] Jiangtao Qiu, Zhangxi Lin, and Yinghong Li. 2015. Predicting customer purchase behavior in the e-commerce context. *Electronic Commerce Research* 15, 4 (2015), 427–452. DOI: https://doi.org/10.1007/s10660-015-9191-6

[70] Ruihong Qiu, Sen Wang, Zhi Chen, Hongzhi Yin, and Zi Huang. 2021. Causalrec: Causal inference for visual debiasing in visually-aware recommendation. In *Proceedings of the 29th ACM international conference on multimedia*. 3844–3852. DOI: https://doi.org/10.1145/3474085.3475266

[71] Hossein A Rahmani, Mohammadmehdi Naghiaei, and Yashar Deldjoo. 2024. A personalized framework for consumer and producer group fairness optimization in recommender systems. *ACM Transactions on Recommender Systems* 2, 3 (2024), 1–24.

[72] P. Ram Mohan Rao, S. Murali Krishna, and A. P. Siva Kumar. 2018. Privacy preservation techniques in big data analytics: A survey. *Journal of Big Data* 5, 1 (2018), 33. DOI: https://doi.org/10.1186/s40537-018-0141-8

[73] Xun Ran, Yong Wang, Leo Yu Zhang, and Jun Ma. 2022. A differentially private matrix factorization based on vector perturbation for recommender system. *Neurocomputing* 483 (2022), 32–41. DOI: https://doi.org/10.1016/j.neucom.2022.01.079

[74] Srishti Rawat, Unnati Tyagi, and Shefali Singhal. 2021. Recommender systems in E-commerce and their challenges. In *Proceedings of the 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*. IEEE, 1598–1601. DOI: https://doi.org/10.1109/ICAC3N53548.2021.9725681

[75] Shaina Raza and Chen Ding. 2022. News recommender system: A review of recent progress, challenges, and opportunities. *Artificial Intelligence Review* 55, 1 (2022), 749–800. DOI: https://doi.org/10.1007/s10462-021-10043-x

[76] F. Ricci, L. Rokach, and B. Shapira (Eds.). 2022. *Recommender Systems Handbook*. Springer US. DOI: https://doi.org/10.1007/978-1-0716-2197-4

[77] Simoni F Rohden and Diully Garcia Zeferino. 2023. Recommendation agents: an analysis of consumers' risk perceptions toward artificial intelligence. *Electronic Commerce Research* 23, 4 (2023), 2035–2050. DOI: https://doi.org/10.1007/s10660-022-09626-9

[78] Markus Schedl, Hamed Zamani, Ching-Wei Chen, Yashar Deldjoo, and Mehdi Elahi. 2018. Current challenges and visions in music recommender systems research. *International Journal of Multimedia Information Retrieval* 7, 2 (2018), 95–116. DOI: https://doi.org/10.1007/s13735-018-0154-2

[79] Nícollas Silva, Diego Carvalho, Adriano C. M. Pereira, Fernando Mourão, and Leonardo Rocha. 2019. The pure cold-start problem: A deep study about how to conquer first-time users in recommendations domains. *Information Systems* 80 (2019), 1–12. DOI: https://doi.org/10.1016/j.is.2018.09.001

[80] Pradeep Kumar Singh, Pijush Kanti Dutta Pramanik, Madhumita Sardar, Anand Nayyar, Mehedi Masud, and Prasenjit Choudhury. 2022. Generating a new shilling attack for recommendation systems. *Computers, Materials & Continua* 71, 2 (2022), 2827–2846. DOI: https://doi.org/10.32604/cmc.2022.020437

[81] Brent Smith and Greg Linden. 2017. Two decades of recommender systems at Amazon. com. *IEEE Internet Computing* 21, 3 (2017), 12–18. DOI: https://doi.org/10.1109/MIC.2017.72

[82] Rama Syamala Sreepada and Bidyut Kr Patra. 2021. Enhancing long tail item recommendation in collaborative filtering: An econophysics-inspired approach. *Electronic Commerce Research and Applications* 49 (2021), 101089. DOI: https://doi.org/10.1016/j.elerap.2021.101089

[83] Yoji Tomita and Tomohiko Yokoyama. 2024. Fair reciprocal recommendation in matching markets. In *Proceedings of the 18th ACM Conference on Recommender Systems*. 209–218.

[84] Duy Thanh Tran and Jun-Ho Huh. 2023. Forecast of seasonal consumption behavior of consumers and privacy-preserving data mining with new S-Apriori algorithm: DT Tran and JH Huh. *The Journal of Supercomputing* 79, 11 (2023), 12691–12736. DOI: https://doi.org/10.1007/s11227-023-05105-6

[85] Andrea C. Tricco, Erin Lillie, Wasifa Zarin, Kelly K. O'Brien, Heather Colquhoun, Danielle Levac, David Moher, Micah DJ Peters, Tanya Horsley, Laura Weeks, et al. 2018. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine* 169, 7 (2018), 467–473. DOI: https://doi.org/10.7326/M18-0850

[86] United Kingdom Government. 2023. A pro-innovation approach to AI regulation. Retrieved August 30, 2024 from https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-papers

[87] Miljan Vučetić and Miroslav Hudec. 2018. A fuzzy query engine for suggesting the products based on conformance and asymmetric conjunction. *Expert Systems with Applications* 101 (2018), 143–158. DOI: https://doi.org/10.1016/j.eswa.2018.01.049

[88] Mengting Wan, Jianmo Ni, Rishabh Misra, and Julian McAuley. 2020. Addressing marketing bias in product recommendations. In *Proceedings of the 13th International Conference on Web Search and Data Mining*. 618–626. DOI: https://doi.org/10.1145/3336191.3371855

[89] Weiquan Wang, Jingjun Xu, and May Wang. 2018. Effects of recommendation neutrality and sponsorship disclosure on trust vs. distrust in online recommendation agents: Moderating role of explanations for organic recommendations. *Management Science* 64, 11 (2018), 5198–5219. DOI: https://doi.org/10.1287/mnsc.2017.2906

[90] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. 2023. A survey on the fairness of recommender systems. *ACM Transactions on Information Systems* 41, 3 (2023), 1–43. DOI: https://doi.org/10.1145/3547333

[91] Zimu Wang, Yue He, Jiashuo Liu, Wenchao Zou, Philip S. Yu, and Peng Cui. 2022. Invariant preference learning for general debiasing in recommendation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1969–1978. DOI: https://doi.org/10.1145/3534678.3539439

[92] Ruoxuan Wei and Hong Shen. 2016. An improved collaborative filtering recommendation algorithm against shilling attacks. In *Proceedings of the 2016 17th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*. IEEE, 330–335. DOI: https://doi.org/10.1109/PDCAT.2016.077

[93] Robyn S Wilson, Adam Zwickle, and Hugh Walpole. 2019. Developing a broadly applicable measure of risk perception. *Risk Analysis* 39, 4 (2019), 777–791. DOI: https://doi.org/10.1111/risa.13207

[94] Bo Xiao and Izak Benbasat. 2015. Designing warning messages for detecting biased online product recommendations: An empirical investigation. *Information Systems Research* 26, 4 (2015), 793–811. DOI: https://doi.org/10.1287/isre.2015.0592

[95] Bo Xiao and Izak Benbasat. 2018. An empirical examination of the influence of biased personalized product recommendations on consumers' decision making outcomes. *Decision Support Systems* 110 (2018), 46–57. DOI: https://doi.org/10.1016/j.dss.2018.03.005

[96] Bo Sophia Xiao and Chee-Wee Tan. 2012. Reducing perceived deceptiveness of e-commerce product recommendation agents: An empirical examination of the relative impact of transparency and verifiability and the moderating role of gender. (2012). Retrieved from https://aisel.aisnet.org/amcis2012/proceedings/HCIStudies/28

[97] Yang Xiao, Qingqi Pei, Lina Yao, and Xianzhi Wang. 2020. Recrisk: An enhanced recommendation model with multi-facet risk control. *Expert Systems with Applications* 158 (2020), 113561. DOI: https://doi.org/10.1016/j.eswa.2020.113561

[98] Yuanbo Xu, Yongjian Yang, En Wang, Fuzhen Zhuang, and Hui Xiong. 2020. Detect professional malicious user with metric learning in recommender systems. *IEEE Transactions on Knowledge and Data Engineering* 34, 9 (2020), 4133–4146. DOI: https://doi.org/10.1109/TKDE.2020.3040618

[99] Chang-Ming Yan and Tzu-Jui Tang. 2011. Applying customer-centered recommendation on an on-line shopping system. In *Proceedings of the 2011 7th International Conference on Natural Computation*, Vol. 4. IEEE, 1993–1997. DOI: https://doi.org/10.1109/ICNC.2011.6022582

[100] Zhihai Yang and Zhongmin Cai. 2017. Detecting abnormal profiles in collaborative filtering recommender systems. *Journal of Intelligent Information Systems* 48, 3 (2017), 499–518. DOI: https://doi.org/10.1007/s10844-016-0424-5

[101] Zhihai Yang, Qindong Sun, and Beibei Zhang. 2018. Evaluating prediction error for anomaly detection by exploiting matrix factorization in rating systems. *IEEE Access* 6 (2018), 50014–50029. DOI: https://doi.org/10.1109/ACCESS.2018.2869271

[102] A. Yashudas, Dinesh Gupta, G. C. Prashant, Amit Dua, Dokhyl AlQahtani, and A. Siva Krishna Reddy. 2024. Deepcardio: Recommendation system for cardiovascular disease prediction using iot network. *IEEE Sensors Journal* 24, 9 (2024), 14539–14547.

[103] Tao Zeng, Xiaohan Fang, Yue Lang, Jiquan Peng, Xi Wu, Shuli Wang, and Jibing Gong. 2021. Fair personalized recommendation through improved matrix factorization by neural networks. In *Proceedings of the 2021 10th International Conference on Networks, Communication and Computing*. 19–24. DOI: https://doi.org/10.1145/3510513.3510517

[104] Fu-guo Zhang and Xu Sheng-hua. 2007. Analysis of trust-based e-commerce recommender systems under recommendation attacks. In *Proceedings of the 1st International Symposium on Data, Privacy, and E-Commerce (ISDPE 2007)*. IEEE, 385–390. DOI: https://doi.org/10.1109/ISDPE.2007.75

Received 4 September 2024; revised 19 October 2025; accepted 19 November 2025