**Published:** 22 November 2025
**Online AM:** 03 October 2025
**Accepted:** 22 September 2025
**Revised:** 13 September 2025
**Received:** 01 December 2024



SURVEY

# A Literature Review of Ethical Considerations in Recommender Systems for User-Generated Content in Human-Computer Interaction

**SHUO NIU**, Clark University, Worcester, MA, United States
**TIANYI LI**, Purdue University System, West Lafayette, IN, United States
**MOHAN CHI**, Purdue University System, West Lafayette, IN, United States

**Open Access Support** provided by:
**Clark University**
**Purdue University System**

# A Literature Review of Ethical Considerations in Recommender Systems for User-Generated Content in Human-Computer Interaction

SHUO NIU, Clark University, Worcester, United States
TIANYI LI, Purdue University System, West Lafayette, United States
MOHAN CHI, Purdue University System, West Lafayette, United States

The design of user-generated content (UGC) platforms poses challenges in comprehensively addressing the ethical dimensions of recommendation algorithms and applying human-centered methods for their evaluation. This article presents a literature review of 97 studies on UGC algorithms (UGCAlgos) that incorporate human factors and user experience considerations to investigate the ethical issues explored in human-computer interaction (HCI) research. Our review identifies key themes in the ethical considerations surrounding UGCAlgos and the user modeling methods employed. We examine how common ethical concerns in recommender systems, such as content appropriateness, privacy, user engagement, transparency, fairness, and diversity, are studied and contextualized within UGC platforms. Furthermore, we summarize how these concerns are addressed through user modeling approaches, including data characterization, user context, user outsmarting, interface design, user beliefs, and community and societal impacts.

CCS Concepts: • **Human-centered computing → HCI theory, concepts and models**; • **Information systems → Content ranking;**

Additional Key Words and Phrases: User-generated content, recommendation, algorithm, literature review

**ACM Reference Format:**
Shuo Niu, Tianyi Li, and Mohan Chi. 2025. A Literature Review of Ethical Considerations in Recommender Systems for User-Generated Content in Human-Computer Interaction. *ACM Trans. Recomm. Syst.* 4, 2, Article 33 (November 2025), 30 pages. https://doi.org/10.1145/3770747

## 1 Introduction

Algorithmic recommendations have become central to social media platforms, enabling the prediction of user interests and the identification of relevant content from large volumes of **User-Generated Content** (**UGC**). Platforms such as X (formerly known as Twitter), Facebook, YouTube, Reddit, and TikTok have intensified their reliance on algorithms for personalization. These UGC platforms predominantly feature non-professional content and support the formation of creator–follower communities [103]. In addition, many UGC platforms employ moderation algorithms to govern the creation of content. The quality of recommendation and moderation algorithms

Authors' Contact Information: Shuo Niu (corresponding author), Clark University, Worcester, Massachusetts, United States; e-mail: shniu@clarku.edu; Tianyi Li, Purdue University System, West Lafayette, Indiana, United States; e-mail: li4251@purdue.edu; Mohan Chi, Purdue University System, West Lafayette, Indiana, United States; e-mail: chi70@purdue.edu.


ACM 2770-6699/2025/11-ART33
https://doi.org/10.1145/3770747

for UGC extends beyond accuracy—these systems influence users socially and emotionally, shape community dynamics, and affect relationships between users and platform governance teams.

While ethical frameworks have been proposed to evaluate **Recommender Systems** (**RecSys**) [120], limited attention has been paid to how the unique context of UGC shapes prevailing ethical considerations. Traditional RecSys evaluations have largely focused on data-centric or algorithm-centric metrics such as accuracy, utility, novelty, and coverage [63, 121]. In UGC recommendation systems, evaluation criteria often center on individual users' perceptions of accuracy [146]. Ethical issues in RecSys research are frequently examined through algorithmic or AI-centric lenses [36, 137], with some studies addressing particular categories of concerns [37] or focusing on specific media types [38]. As a result, there remains a lack of human-centered understanding of how ethical concerns shape diverse aspects of user interaction with **UGC algorithms** (**UGCAlgos**) and extend beyond content consumers to include content creators, platform governance structures, and broader societal implications.

In addition to framing the ethical considerations of UGCAlgos, current methods for evaluating user experience in UGC recommendation systems are primarily categorized as offline analytics, user studies, and online experiments [48, 121, 146]. However, there remains limited understanding of which human-centered methods have been employed to address various ethical considerations in UGCAlgos.

In this literature review, we adopt a **Human-Computer Interaction** (**HCI**) perspective to examine existing HCI research on UGC recommendation systems. Drawing on standard definitions of RecSys [14, 92], we define UGCAlgos as *"algorithmic systems that attempt to recommend or moderate UGC by associating information about users' preferences, knowledge about the content, and social information."* We conduct a descriptive analysis of 97 papers published in HCI venues that pertain to UGCAlgos [107]. Our goal is to summarize how ethical concerns are uniquely shaped within the UGC ecosystem, while also identifying common HCI techniques for modeling user behavior to address ethical challenges. Specifically, we address the following two research questions:

— RQ1: How does the HCI literature describe the ethical challenges of RecSys for UGC, and what evaluation criteria should be considered to mitigate these ethical concerns?
— RQ2: How does the HCI literature model users' preferences for more ethical RecSys for UGC, and how do these user modeling techniques address different ethical concerns?

We ground our review in the framework proposed by Milano et al. [100], which outlines ethical challenges related to *content*, *privacy*, *autonomy and personal identity*, *opacity*, *fairness*, and *social effects* in RecSys (see Table 1, left). This framework offers a foundational understanding of common ethical challenges in RecSys. In this review, we contextualize it for UGC platforms by examining how HCI studies address UGCAlgo ethics (see Table 1 for our definitions). We also identify common human-centered methods used to understand and address these concerns (Table 2). This review contributes by identifying six major themes through which the ethical considerations of UGCAlgos are examined in HCI research: (1) mitigating inappropriate UGC through understanding human behavior, (2) protecting user privacy in interface design, (3) capturing true user interest and enhancing healthy engagement, (4) providing transparency beyond explaining recommendation algorithms, (5) ensuring fairness in visibility, moderation, and user interaction, and (6) incorporating diversity evaluation metrics, social-based strategies, and interactive diversification.

## 2 Background

### 2.1 User-Generated Content and Algorithmic Curation

UGC refers to information or media voluntarily contributed by individuals for public use [78]. Naab et al. identified three key criteria that define UGC: (1) it involves substantial individual contributions,

Table 1. The Ethical Consideration Subthemes in the Second Column are Defined by Milano's Framework [100], while Those in the Third Column are Reinterpreted in this Review to Reflect the Specific Context of UGCAlgos

| **Definition in Milano's Framework [100]** | **Definition in the UGC Context** |
|---|---|
| *Inappropriate content* in [100] refers to recommending content that conflicts with users' ethical or cultural values, violate privacy, or expose users to biased, harmful, or manipulative material. | In this review, **Adversarial Content Dynamics** refers to studies that investigate the dynamics of inappropriate content production and its associated internet subculture communities, their impacts on users, and the evolving implications for platform moderation policies. |
| *Privacy* in [100] is defined as the protection of users from unauthorized data collection, storage, inference, and profiling without informed consent. | In this review, **Privacy Configuration** refers to studies that evaluate how user requirements for privacy settings impact user interactions and RS interface design. |
| *Autonomy and Personal Identity* in [100] refers to the ethical challenge where these systems can encroach on users' autonomy by nudging them in specific directions, attempting to "addict" them to certain types of content. | In this review, **Interest Modeling and User Reflection** encompasses studies using various sources and design techniques that enable UGCAlgos to more accurately capture users' genuine interests while mitigating the effects of filter bubbles. It also includes approaches that prompt users to reflect on their engagement with UGC. |
| *Opacity* in [100] concerns improving transparency in how users are classified and modeled to mitigate algorithmic bias. | In this review, **Transparency Structure** encompasses transparency needs across UGCAlgos, including consumers' demand for governance-related explanations of recommendations, fair mechanisms for creators to contest algorithmic decisions, and transparent controls that grant agency over both content delivery and reception. |
| *Fairness* [100] is defined as the effort to prevent algorithmic outcomes from reproducing or reinforcing social biases, either by mitigating disparities between user groups or by balancing competing user interests. | In this review, **Socio-Algorithmic Fairness** evaluates UGCAlgos in terms of the representation of diverse creator groups, the equitable allocation of platform resources, the fairness of governance in moderating UGC, and the extent to which consumers have effective control over algorithmic outcomes. |
| The *Social Effects* in [100] refers to preventing RecSys from reinforcing social biases or enabling manipulation, by promoting diversity, reducing informational segregation, and balancing individual preferences with democratic values. | In this review, **Diversification** encompasses studies that evaluate the breadth of recommendations through algorithmic diversification, incorporate social dynamics to enhance diversity, and explore interactive approaches to support content discovery. |

emphasizing the personal nature of the content; (2) it must be publicly accessible; and (3) it is created outside professional contexts and practices [103]. UGC plays a vital role in self-expression and engagement, encompassing a variety of formats such as videos, text, audio, images, blogs, and networking information [102]. Social media platforms have empowered non-professionals to create and share content, significantly increasing the volume of online content. These platforms foster unique cultures shaped by user interactions and contributions. UGC enables individuals to express personal opinions, share diverse perspectives, and distribute decentralized content, which forms dynamic and evolving repository of information and creativity [117].

Algorithms have fundamentally transformed the landscape of UGC platforms [91]. Leading platforms such as Facebook [11, 46, 47], X/Twitter [10, 19, 60], YouTube [34, 138, 142], and TikTok [73, 83, 133] have developed *curation algorithms* that sort, filter, and disseminate content feeds. These algorithms are designed to detect emerging trends on social media and recommend content, that is,

Table 2. Subthemes of User Modeling Methods Synthesized from HCI Research on RecSys, Identified, and Defined Through Thematic Analysis of the Reviewed Papers

| Subtheme | Definition |
|---|---|
| **User and Content Characterization** | Analyze UGC based on content attributes such as topic, hashtags, and modality, along with metadata attributes like views, likes, posting location, and time; or characterize user profiles using account attributes and interaction behaviors. |
| **Contextual Impact on Consumers** | Investigate how UGC recommendations impact individual intentions, emotions, and feelings. |
| **Creators' and Consumers' Folk Thoery** | Assess creators' informal beliefs about platform governance and consumers' informal understandings of UGCAlgo mechanisms and quality. |
| **Creator Algorithm and Policy Outsmart** | Examine the strategies creators use to counteract, circumvent, or outsmart UGCAlgos and governance policies. |
| **Recommendation Visualization and Control** | Develop novel interfaces and visualizations that contrast with traditional ranking-based recommendations. |
| **Community and Societal Impact** | Evaluate the socio-cultural, socio-economic, and socio-technical impacts of UGCAlgos, or study how they affect or should be designed to support underrepresented and vulnerable groups. |

personally relevant and engaging [117]. They also provide suggestions for friends or professional connections [61, 65, 141].

This environment has fostered grassroots communities of content creators who can rapidly gain visibility and influence by reaching relevant users [90, 139]. UGCAlgos must accommodate non-professional content production, filter inappropriate posts [32], and address broader social and community implications [71]. To sustain a safe and positive space, algorithmic content moderation has become a central mechanism of platform governance [57]. As a consequence, creators increasingly seek to understand and adapt to algorithmic logic, while consumers desire more control over how content is curated and presented [3]. Unlike traditional RecSys that recommend single-category data, UGCAlgos must navigate complex social and community dynamics. However, despite their significant influence, few reviews examine how human-centered research defines the responsibilities and ethical implications of UGCAlgos, including their individual and societal impacts and the development of user-centered evaluation criteria.

### 2.2 Recommendation Algorithms and Evaluation

Recommendation systems are typically defined as programs that suggest suitable items (products or services) to users by predicting their interests based on information about the items, the users, and their social interactions [14, 92, 108]. Moderation algorithms are designed to block or flag UGC based on predictive techniques, thereby producing a governance outcome [57]. These systems utilize various algorithmic and AI methods [67]. Content-based, collaborative filtering, and hybrid approaches have been widely adopted in the development of recommendation algorithms [108].

Research has emphasized that the evaluation of RecSys should extend beyond accuracy-based metrics [48]. Common evaluation methods include offline analytics, laboratory-based user studies, and real-world online testing [48]. Kuanr and Mohapatra reviewed both data-centric metrics (e.g., accuracy, MAE, and coverage) and user-centric evaluation approaches [79]. User-focused frameworks, such as ResQue, highlight dimensions like novelty, diversity, explanation, and transparency [110].

Silveira et al.'s framework emphasizes utility, unexpectedness, and serendipity [121]. Additional models, including those by Olmo and Gaudioso [63] and Zangerle and Bauer's FEVR framework [146], underscore the importance of evaluating filtering processes, item presentation, and broader evaluation principles.

The aforementioned evaluation methods of RecSys focus on metrics and user satisfaction; however, macro-level evaluations of UGC recommendation systems must also consider their emotional, ethical, social, and economic impacts [132]. While human-centered research has highlighted sociotechnical impacts, there is limited synthesis of how HCI studies tackle ethical challenges at individual, community, and societal levels, which is essential for designing ethically sound UGC recommendation algorithms.

### 2.3 The Ethical Challenges of RecSys for User-Generated Content

Algorithms for recommending and moderating UGC present numerous ethical challenges for individuals and communities [21, 76]. Milano et al. identified six primary areas of ethical concern: inappropriate content, privacy violations, autonomy and personal identity, opacity, fairness, and social effects [100]. Although this categorization outlines the overarching ethical challenges of RecSys, these challenges pose unique implications for UGC creators and users, requiring additional considerations in the design of UGCalgos. For instance, the fairness concern highlighted in [100] primarily addresses issues such as unfair collaborative filtering or biases in representing user interests. However, fairness in the context of UGCalgos encompasses broader issues, including whether content creators receive appropriate visibility [100] and whether fair mechanisms exist for appealing moderation decisions [130]. Similarly, the opacity discussed in pertains to how users are modeled and whether recommendations are sufficiently explained. Beyond this, users' expectations for UGC recommendation systems also involve their confidence that these algorithms align with their sense of control over the recommendation outputs [1].

Algorithm-centric research on ethical issues in RecSys has primarily focused on data modeling, algorithm design, and user settings to address specific concerns within particular domains. Shani and Gunawardana [120] emphasized user experience properties in RecSys and recommended consistent data, parameters, and evaluation. In music recommendation, Deldjoo et al. [38] advocated for enhancing data diversity and algorithmic context-awareness. Privacy-oriented research has addressed ethical data collection and algorithmic support for user privacy settings [55, 87, 96, 150]. To improve transparency, some studies have employed natural-language summaries [7]. Fairness remains a central issue, with methods such as excluding user reputations from ranking algorithms proposed to reduce bias and manipulation [113, 147]. Content diversity has been improved using techniques like knowledge graphs [6, 49].

Beyond algorithmic solutions, researchers have stressed the importance of addressing the ethical issues of RecSys by integrating the rights of different stakeholders alongside utility metrics [99]. Broader ethical concerns have called for human-centered and society-centered qualitative evaluation methods [72]. Reviews of inappropriate content in RecSys emphasize the need for interdisciplinary approaches that combine computational, social science, and ethical perspectives while accounting for diverse stakeholder viewpoints [55, 135]. Deldjoo et al. [37] advocate expanding fairness assessments beyond algorithmic metrics to include human judgment and societal impact. Algorithmic solutions often overlook the multifaceted nature of ethical concerns and their varying impacts across stakeholders on UGC platforms. This review responds to calls for an interdisciplinary understanding of how ethical considerations affect different stakeholder groups within the unique social context of UGC environments.

Understanding how HCI research models users' responses to ethical challenges is essential for developing evaluation criteria for UGCAlgos [76, 110]. This includes not only assessing perceived

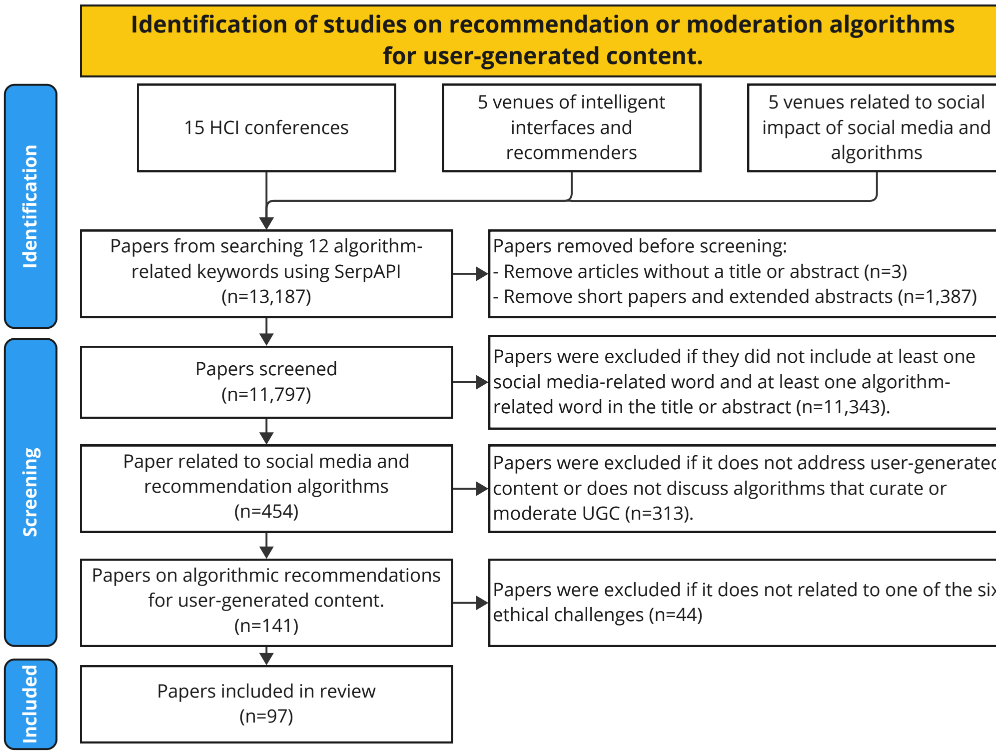


Fig. 1. PRISMA diagram illustrating the steps of paper identification, screening, and inclusion.

recommendation quality but also examining how users interpret opaque algorithmic decisions through informal "folk theories" [40, 73, 144]. Yet, there is still a lack of comprehensive insight into the methods HCI researchers use to model users' reactions to the ethical dimensions of UGCAlgos.

## 3 Methods

### 3.1 Paper Identification

Guided by Naab et al.'s definition of UGC [103] and Milano's ethical framework for RecSys [100], our review focuses on UGCAlgo-related papers in HCI that address the six ethical dimensions identified for general RecSys (Table 1). Through a descriptive review process—including paper searching, screening, and classification (Figure 1 shows the overall paper search and screening steps)—we aim to uncover patterns and gaps in the literature related to existing theories [107]. Our search specifically focuses on identifying a representative sample of HCI works that address the ethical challenges associated with UGC recommendation and moderation algorithms, as well as the human-centered user modeling methods used to study and address these challenges [107]. We acknowledge the existence of research on RecSys ethics in algorithm-focused and AI-focused venues such as SIGIR, TOIS, and WWW. However, our aim is to highlight the distinct perspectives offered by HCI research within the UGC context and to identify human-centered methodologies for addressing ethical challenges.

We selected 25 venues where papers on user experiences on social media are usually published, sourcing from mixed places like Google Scholar[1] and SIGCHI website[2]. The details of these

[1]https://scholar.google.es/citations?view_op=top_venues&hl=en&vq=eng_humancomputerinteraction
[2]https://sigchi.org/conferences/upcoming/

conferences and journals are in Appendix A. The venue search was focused on HCI venues where papers on interactions with social media or recommendation interfaces could be published. We selected 15 core HCI conferences (e.g., CHI, CSCW, UIST, and DIS). We also included five venues on the themes of intelligent interfaces and recommenders (RecSys, IUI, TiiS, CHIIR, and HT) and five venues related to social impact of social media and algorithms (COMPASS, UAIS, NMS, IJHCI, and FAccT). These venues cover HCI, social computing, social media research, and recommendation systems, offering a representative sample for an HCI-centered understanding of UGCAlgo research.

To compile a list of search keys, we first identified 10 sample papers that match our research goal and examined their key words sections to identify words related to algorithmic recommendation. The resulting phrase list includes terms such as "algorithm recommendation," "algorithmic recommendation," "recommendation algorithm," "RecSys," "algorithmic curation," "algorithmic experience," "algorithmic moderation," "algorithm content," "algorithmic content," "algorithm folk theory," "algorithm auditing," and "algorithm evaluation." This process resulted in a review pool of 11,797 papers. Duplicate papers were merged based on their DOI, and short papers and extended abstracts were excluded. We did not include UGC concepts in the initial search due to the large number of platform names that needed to be considered. The articles related to UGC were filtered using a word embedding search, as described below.

## 3.2 Paper Screening

Given the vast number of returned papers about algorithms, we adopted a two-stage filtering process—first using a computer program for coarse filtering, followed by manual annotation and discussion—to identify papers that are closely relevant to the study of ethical challenges and user modeling on UGC platforms.

*3.2.1 Programmatic Screening.* In the programmatic filtering stage, we used a script to check whether the title or abstract contained at least one social media and one algorithm-related keyword. Text was lowercased, tokenized, and stemmed for accurate matching. Social media keywords were drawn from platform names listed on Wikipedia[3] and a set of common social media terms. This Wikipedia list includes platforms with a minimum of 100 million active users, such as "Facebook," "YouTube," "WhatsApp," "Instagram," "TikTok," "Messenger," and many others (see Figure 2 Right for the article distribution). General social media terms include "UGC," "social media," "streaming," "blog," "podcast," "forum," "networking," and "video sharing." Moreover, the algorithm-related keywords used to filter the articles are "algorithm," "recommend," "curation," and "moderation" These algorithmic terms were selected to ensure that the articles discuss algorithmic processes. After this filtering process, we reduced the article pool to 454. It is worth noting that some social media platforms do not exclusively feature UGC (e.g., Messenger), while others combine public and private content (e.g., WeChat). Nevertheless, we included such platforms in the coarse filtering to avoid excluding relevant papers from cross-platform studies.

*3.2.2 Manual Screening Procedure.* During the manual annotation by the researchers, three authors of this article independently reviewed the 454 papers to determine whether each should be included in the review. For a paper to be included, it had to simultaneously meet three criteria:

Criteria 1. The article must address UGC as defined in [103]. The targeted social media content should have an element of personal contribution, be shared publicly so that everyone can access it, and not be primarily produced by professionals or for commercial purposes. Therefore, chat content on apps like WhatsApp (not publicly shared), music and movie

[3]https://en.wikipedia.org/wiki/List_of_social_platforms_with_at_least_100_million_active_users

streaming on platforms like Spotify and Netflix (only professional production), and room/ride-sharing information on apps like Airbnb or Uber (mostly for commercial services) were not considered UGC.

Criteria 2. The article must discuss algorithms that curate or moderate UGC, and it must include considerations of user and social dynamics, such as user interaction data, user profiles, or community measurements. This criterion is satisfied if the article addresses algorithm lifecycles (e.g., design, development, evaluation, and interfacing) or conducts empirical analyses of algorithmic influence on users, communities, and society. A few papers focused on algorithmic contributions (e.g., [42, 60]) were included because they consider human factors, social relationships, or community structure in algorithm design. However, papers solely focused on social media interface design, purely human activity, content editing algorithms, or attribute-detection algorithms without addressing recommendation or moderation impacts were excluded.

Criteria 3. The article must address at least one ethical dimension as outlined in [100]. We evaluated each paper for relevance to six key ethical concerns: appropriateness, privacy, interest and engagement, transparency, fairness, and diversity. Final alignment with these ethical dimensions was determined through author discussions.

During the manual annotation phase, we first applied Criteria 1 and 2 to filter the initial set of 454 papers. Three authors independently evaluated each paper, indicating whether each criterion was satisfied. This process resulted in 141 papers retained for further review.

To meet Criteria 3 and identify papers pertinent to the research questions, the 141 papers were divided into four batches ($N = 30, 30, 50, 31$). The development of themes with the first 30 papers followed Milano's framework [100] for ethical dimensions, whereas user modeling themes were derived through thematic analysis, as outlined in [131]. In the first round, three authors independently reviewed 30 papers, leaving notes about their connections to ethical dimensions and user modeling methods. Titles and notes were converted into digital cards (Miro.com) and we grouped the cards using an affinity diagram. This round produced initial themes, definitions, and example papers. For ethical considerations, papers were grouped according to the original dimensions in [100], and each dimension was compared and contrasted between general RecSys [100] and the UGC context. The the second set of 30 papers was used to validate and refine the themes. In cases of disagreement, the authors re-examined the article and discussed it to reach a consensus, either by clarifying the sections where the ethical concern was mentioned in the article or by revising the definition of the ethical subthemes. If a paper relates to multiple ethical considerations in different sections, we categorize it under all relevant ethical considerations. In the final round, the authors independently labeled the 31 papers using the ethical consideration and user modeling subthemes. Because each paper could be assigned multiple labels, Krippendorff's alpha was used to assess inter-rater reliability. The results, presented in Table 3, indicate that most themes achieved tentative to substantial levels of agreement. Most subthemes were well-aligned, except "Interest Modeling and User Reflection," which had lower agreement due to inconsistent interpretation. All discrepancies were discussed and resolved to finalize the annotations.

For RQ1, we assess how the subthemes require updated definitions from [100], given that the unique dynamics and ethical considerations of UGC necessitate a re-examination of the original framework. Table 1 compares how the dimensions proposed by Milano were adopted or modified. For RQ2, we identified six subthemes that describe how HCI research involves users in the design or evaluation of algorithms (Table 2). Papers that did not align with any dimension were excluded based on Criterion 3, resulting in the removal of 44 papers. Ultimately, 97 papers remained in the final review pool.

Table 3. The Inter-rater Agreement Score for all Subthemes Was Calculated Using Krippendorff's Alpha

| | | | | | |
|---|---|---|---|---|---|
| Adversarial Content Dynamics | Privacy Configuration | Interest Modeling and User Reflection | Transparency Structure | Socio-Algorithmic Fairness | Diversification |
| 0.816 | 1 | 0.524 | 0.743 | 0.762 | 0.850 |
| User and Content Characterization | Contextual Impact on Consumers | Creators' and Consumers' Folk Thoery | Creator Algorithm and Policy Outsmart | Recommendation Visualization and Control | Community and Societal Impact |
| 0.741 | 0.658 | 0.868 | 1 | 0.657 | 0.752 |

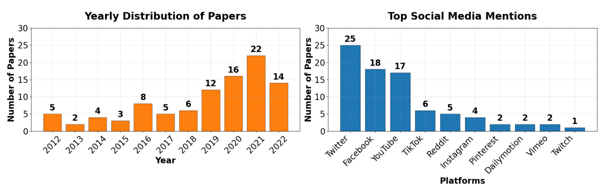


Fig. 2. Left: Yearly distribution of papers. Right: The number of papers mentioning the top 10 social media names in their titles or abstracts.

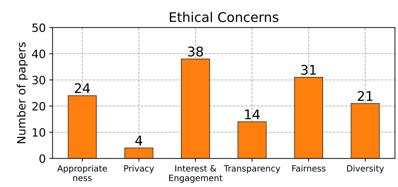


Fig. 3. Number of papers in each ethical concern subtheme.

## 4 Results

Our analysis shows an increase in the number of papers that addresses UGCAlgo over the past decade (Figure 2). Notably, there's a growing focus on the ethical concerns within UGC platforms, as well as strategies to enhance user agency in addressing those issues. CSCW (24), NMS (12), and RecSys (11) are the top venues where these papers were published. One paper each from DIS, IJHCI, and TOCHI was identified. Below, we first discuss the ethical concerns associated with UGCAlgos that were raised from those papers. We then summarize the HCI perspectives on addressing these ethical concerns through user modeling.

### 4.1 RQ1: Ethical Consideration of UGC Recommendation and Moderation Algorithms

In this section, we present how the six challenges identified in Milano's framework [100] manifest in UGCAlgos from an HCI perspective (see Table 4 for an overview and Figure 3 for the number of papers associated with each subtheme).

*4.1.1 Adversarial Content Dynamics.* HCI research has examined inappropriate content primarily in relation to specific problematic online behaviors or vulnerable user groups, thereby offering insights into the dynamic nature of inappropriate content that conflicts with users' ethical or cultural values. Our analysis identified 24 papers that investigated three main types of content that UGCAlgos are tasked with moderating and mitigating.

*Misinformation and manipulative practices.* Understanding specific forms of misinformation and manipulative practices on UGC platforms requires content characterization, analysis of manipulative user behaviors, evaluation of user impacts, and investigation of moderation strategies. HCI research has emphasized the nature of such content and the behaviors that facilitate its spread, including bot-driven misinformation [24, 127], strategies for inflating popularity metrics such as clickbait and pod activities [17, 85, 136], and methods for circumventing moderation efforts [53, 139]. Additionally, through algorithm auditing and user simulation, studies have found that query intent and personalization based on watch history significantly influence users' exposure to misinformation [66, 123, 128]. In response, some studies propose frameworks for explaining fake

Table 4. An Overview of Ethical Considerations, Including their Subthemes and Example Papers

| **Ethical Consideration** | **Subtheme** | **Paper** |
|---|---|---|
| Adversarial Content Dynamics | Misinformation and manipulative content | [127], [24], [17], [136], [85], [53], [139], [66], [128], [123], [30], [70], [93] |
| | Hate speech and harassment | [115], [81], [106], [51], [97], [35], [20], [45], [124], [56] |
| | Health-harming content | [26], [53], [106], [97], [20] |
| Privacy Configuration | - | [93], [119], [86], [46] |
| Interest Modeling and User Reflection | Improving how user preferences are captured and understood | [101], [84], [46], [140], [143], [119], [114], [149], [64], [29], [2], [148], [9], [142], [145], [134], [41], [11], [90], [89], [52], [59], [27], [31], [5], [133] |
| | Ensuring fair representation of diverse content and interests. | [88], [4], [98], [109], [25], [12], [69], [31] |
| | Fostering user reflection and critical thinking | [23], [125], [16], [46], [105], [62] |
| Transparency Structure | Explanatory transparency | [112], [30], [16], [86], [43], [44], [111], [139], [3], [80] |
| | Contestability transparency | [130], [129], [94], [86] |
| | Control transparency | [94], [46], [86], [16], [112], [3] |
| Socio-Algorithmic Fairness | Representational fairness | [73], [122], [39], [31], [116], [54], [15], [33], [50], [19], [25], [22], [59] |
| | Allocative fairness | [13], [18], [39], [33], [94], [74], [43], [95] |
| | Governance fairness | [56], [129], [122], [35], [45], [82], [130], [43], [95] |
| | Interactional fairness | [88], [111], [44], [19] |
| Diversification | Algorithmic diversification | [12], [8], [66], [128], [88], [104], [42], [75], [109] |
| | Social diversity | [25], [60], [142], [118], [12], [114], [51], [59] |
| | Interactive exploration | [126], [28], [58], [50], [59], [12], [80] |

news detection outcomes [30] or evaluating content credibility [70], while others examine how misleading content prompts user disengagement from platforms [93].

*Hate speech and harassment.* This line of HCI research aims to understand how UGCAlgos facilitate extremist social activities, inform moderation policy needs, and shape users' perceptions of moderation. Algorithms and UGC platforms have been found to facilitate the formation of online spaces where radical or extremist groups can connect [35, 51, 81, 97, 106, 115]. HCI researchers have emphasized the need for UGCAlgo policies that protect vulnerable groups such as youth [20] and immigrants [45], and have examined users' perceptions of abusive and discriminatory content [124] as well as the outcomes of moderation efforts [56].

*Health-harming content.* A notable theme in harmful content investigated by HCI researchers involves health-related topics, focusing on the nature of the content, the creator communities that produce it, and the techniques used to evade moderation. Studies have investigated health-harming content such as pro–eating disorder material [26], including how such content bypasses platform moderation [53]. HCI researchers have also explored communities centered around sensitive and potentially harmful sexual discourse, such as misogynistic and incel groups [97, 106]. While this content is highly sensitive, research has highlighted the challenges faced by social service providers in moderating it to protect youth [20].

*4.1.2 Privacy Configuration.* HCI studies on privacy configuration evaluate users' demands for privacy settings and examine how interface design can provide greater privacy control. Four papers in our review address users' privacy needs. Privacy concerns have been identified as a key reason why some users avoid certain UGC platforms [93]. HCI researchers have recommended anonymizing user interaction data in the design of UGC recommendation interfaces [119], as well as equipping users with tools to understand and control how their personal information is used in generating recommendations [46, 86].

*4.1.3 Interest Modeling and User Reflection.* A central topic in HCI research on human–RecSys interaction is understanding how user interests are captured from various sources and developing techniques that encourage users to reflect on their own interests. Our review identified 38 papers related to this subtheme, making it the most prevalent ethical consideration in our dataset.

*Improving how user preferences are captured and understood.* HCI studies have examined how UGCAlgo designs can better reflect users' authentic selves and prioritize their values, rather than simply maximizing engagement or encouraging addictive behaviors [46, 84, 101]. User engagement must be balanced with goals such as diversity and novelty [140]. To support this, human-centered techniques leverage implicit signals from natural interactions with UGC—such as clicks [119, 143] and browsing history [2, 29, 64, 114, 148, 149]—to model user preferences. Social network information has also been used to enhance personalization and foster interaction [9, 11, 41, 90, 134, 142, 145].

Other approaches incorporate data from multiple platforms [52, 89] and geolocation [59]. Some studies highlight the need for differentiated strategies across user groups [27, 31], and others stress the importance of tailoring recommendations to specific content types (e.g., GIFs [5]). Beyond preference modeling, research also shows that different forms of engagement can shape user behaviors, such as purchase intentions [133].

*Ensuring fair representation of diverse content and interests.* The HCI community has examined challenges such as filter bubbles and over-personalization. Both user preferences and UGCAlgos contribute to the formation of various types of filter bubbles [88]. Techniques for generating sufficiently diverse recommendations have been explored through algorithmic approaches [4, 25, 98, 109], as well as by leveraging users' social connections [12, 69]. Ensuring fair representation is particularly critical for UGC users from marginalized groups [31].

*Fostering user reflection and critical thinking.* Many HCI studies address this ethical concern by designing interfaces that prompt users to reflect on the relevance of content and provide explanations for why specific recommendations are made [16, 23, 46, 125]. To further support user autonomy, researchers emphasize the importance of collecting both explicit and implicit feedback on recommendations. Explicit feedback, often obtained through conscious actions such as likes, comments, and views, can be complemented by implicit signals such as facial expressions [105] or historical interactions with prior recommendations [62]. Collecting and integrating such implicit feedback not only enhances users' self-awareness of their UGC consumption but also enables iterative improvements to the system's responsiveness and transparency [23, 62, 105].

*4.1.4 Transparency Structure.* As highlighted in 14 research papers, HCI research has delineated how to not only explain recommendation results but also establish governance structures, thereby enabling creators to contest decisions and allowing content consumers greater control over what they consume.

*Explanatory transparency.* This aspect underscores the importance of UGCAlgos providing clear explanations of how they operate, make decisions, and rank or censor content [16, 30, 43, 44, 86, 111, 112]. Transparency can be supported not only through explanation but also through the design of intuitive, user-friendly interfaces that help users interpret algorithmic decisions. Content creators' understanding of UGCAlgos significantly influences their behavior on these platforms [3, 139]. Enhancing explanatory transparency and raising content consumers' awareness of algorithmic processes—including inherent biases and filtering mechanisms—are critical for fostering trust and promoting informed engagement [3, 80, 111].

*Contestability transparency.* HCI research has been advocating the need to make algorithms contestable—providing creators more power to challenge or appeal decisions made by UGCAlgos, such as content deletion or demonetization actions [129, 130]. Contestability in the criteria and processes used for moderating certain UGC is essential to ensure that content creators feel they are being treated fairly on the platform [86, 94].

*Control transparency.* The control of UGCAlgos must also ensure transparency about how users' intended actions are interpreted and executed by the system [46, 94]. Such transparency helps both creators and consumers understand the broader implications of UGCAlgos on their interactions with content, including the economic, social, and trust-related consequences of algorithmic decisions on the platform [46, 86, 94].

This aspect is also closely tied to the interactive exploration methods mentioned earlier, as enabling users to view and adjust recommendation results through direct manipulation fosters a deeper understanding of algorithmic processes [16, 112]. HCI researchers have also explored methods for eliciting users' needs and goals with algorithm outcomes, such as directly asking content creators to specify their expected target audience [46] and evaluating consumers' "algorithm experience" to identify gaps in control and transparency [3].

*4.1.5 Socio-Algorithmic Fairness.* Our review identified 31 papers that explore socio-algorithmic fairness. HCI research views fairness on UGC platforms as a socio-technical concern uniquely shaped by the dual role of users as both content creators and consumers. Fairness issues are especially salient for users with minority identities. It extends beyond recommendations to include whether creative work is fairly reimbursed and whether governance decisions are equitable. While some studies focus on traditional notions of fairness—such as selecting a balanced set of topics from candidate UGC [80], addressing missing data [114], and developing metrics for evaluating fairness [68]—others emphasize the iterative, social, and governance-related dimensions of fairness.

*Representational fairness.* Recommendation algorithms must carefully manage how content from different groups is represented and moderated. UGCAlgos should avoid discrimination based on identity attributes such as race, gender, disability, sexual orientation, and other social categories [15, 31, 39, 54, 73, 116, 122]. Beyond recommendations, moderation algorithms must ensure that creators are not disadvantaged due to underrepresentation or disproportionately penalized compared to other groups [54, 73]. Creators' perceptions of how identity shapes algorithmic processes can significantly influence their behavior on UGC platforms [33, 54, 73]. Moreover, recommendation systems should aim to balance diverse political perspectives [19, 22, 25, 50] and amplify content from individuals across different geographic regions [59].

*Allocative fairness.* Content creators demand allocative fairness, as they depend on UGCalgos to maximize the reach and engagement in order to be rewarded for their creative labor. However, content from creators belonging to marginalized groups may be less likely to be recommended compared to content from more popular or mainstream creators [13, 18, 33, 39].

While UGC platforms increasingly employ algorithmic content moderation to detect and penalize policy violations—through measures such as shadow banning or demonetization—these actions are often perceived as reinforcing stereotypes and limiting creators' ability to generate income [94]. HCI research has highlighted these negative experiences and fairness concerns regarding the distribution of visibility, resources, and benefits across different creator groups [43, 74, 94]. These UGCAlgos are frequently perceived as biased and accused of exacerbating inequities, particularly for minority groups such as LGBTQ+ creators [39, 43, 74, 95].

*Governance fairness.* A critical area identified in the HCI literature involves the governance policies of UGCAlgos and the lack of transparency surrounding how these algorithms make decisions. Algorithmic censorship is often perceived as a threat to free expression and as a mechanism that suppresses diverse viewpoints [56, 122, 129]. Platform policies and algorithmic practices – such as inconsistencies in moderation outcomes and the degree to which users can participate in governance—can significantly shape creators' perceptions of fairness [35, 45, 82]. HCI studies underscore the importance of fair appeal processes [129, 130], consistency in algorithmic decisions [43, 95], and mechanisms to prevent discrimination [56].

*Interactional fairness.* HCI research has also examined users' expectations of fairness in how their interactions are interpreted by UGCAlgos and in the extent to which they can control algorithmic outputs. Interactive solutions have been proposed to help ensure that content consumers receive fair and balanced recommendations [88]. Content creators, in particular, often report unmet expectations or a lack of control when interacting with UGCAlgos [44, 111]. Content consumers similarly seek clarity regarding who controls algorithmic decisions and whose values are embedded in those processes [19]. These concerns underscore the need for more transparent explanations of algorithmic behavior and greater user control over algorithmic recommendation [44, 111].

*4.1.6 Diversification.* HCI research examines diversification from the perspective of employing diverse mechanisms to broaden what consumers see in posts. Twenty-one papers in our review addressed this dimension.

*Algorithmic diversification.* Both content creators and consumers benefit from the deliberate diversification of recommendations, tailored to various criteria and contexts [12]. Auditing studies have shown that UGCAlgos can produce homogeneous misinformation content and contribute to the formation of echo chambers [8, 66, 128]. To address this, recommendation evaluation criteria should be developed to assess the extent to which UGCAlgos expose users to diverse perspectives while mitigating polarization [42, 88, 104]. HCI research further underscores the importance of adapting diversification criteria to specific content types [75] and employing heuristic methods to optimize the selection of recommended items [109].

*Social diversity.* HCI research underscores the importance of leveraging social connections and networks to diversify recommendations [25, 60]. This includes promoting cross-network collaboration to uncover user interests [142] and fostering new social ties on UGC platforms [12, 118]. However, users also tend to revisit content from the same creators, reinforcing existing preferences [114]. UGCAlgos should avoid relying solely on user upvotes, as this can amplify polarized content from non-diverse sources [51]. Furthermore, research cautions that online discourse is often dominated by voices from more developed regions, limiting geographic diversity [59].

*Interactive exploration.* HCI research has explored recommendation interfaces that enable users to interactively explore diverse information, moving beyond the traditional linear presentation of

Table 5. An Overview of user Modeling Method and Example Papers

| User Modeling Subtheme | Paper |
|---|---|
| User and Content Characterization | [75], [9], [145], [80], [88], [114], [127], [2], [98], [58], [62], [17], [124], [24], [16], [59], [52], [26], [118], [115], [136], [86], [15], [123], [148], [41], [28], [70], [27], [11], [5], [81], [69], [66], [46], [4], [43], [53], [85] |
| Contextual Impact on Consumers | [130], [133], [125], [12], [93], [111] |
| Creators' and Consumers' Folk Thoery | [139], [94], [74], [13], [19], [73], [56], [30], [112], [44], [116], [95], [93], [111], [82], [129], [122], [46], [31], [39], [33], [22] |
| Creator Algorithm and Policy Outsmart | [139], [94], [13], [19], [73], [136], [116], [82], [122], [31], [39], [53], [33], [51] |
| Rec. Visualization and Control | [126], [58], [50], [104], [124], [23], [16], [59], [3], [44], [125], [12], [86], [46] |
| Community and Societal Impact | [130], [94], [74], [20], [68], [73], [35], [59], [42], [116], [97], [54], [60], [122], [31], [45], [39], [18], [51] |

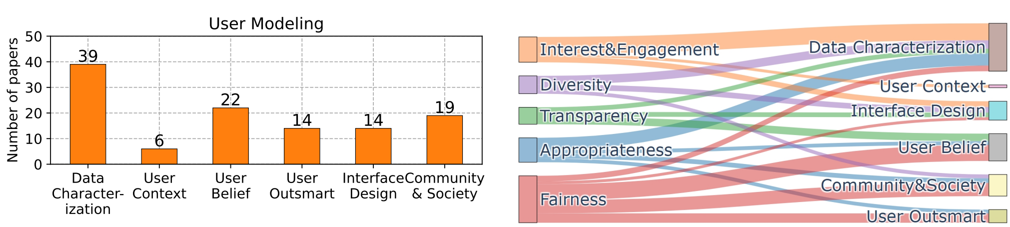


Fig. 4. Left: Number of papers in each user modeling method. Right: Association between ethical dimensions and user modeling methods. Each flow represents the number of papers that simultaneously mention an ethical issue and a user modeling method. The figure includes only associations mentioned by at least three papers. The width of each strip is proportional to the number of papers.

recommended items [28, 126]. Some approaches visualize relationships among recommended items based on criteria such as topical distance or opposing perspectives on sensitive issues [50, 58, 59]. Others facilitate the exchange of personal recommendation lists between users to expose alternative preferences and interests [12]. Interactive interfaces also play a key role in empowering users to identify and mitigate potential biases in UGCAlgos [80], while encouraging critical reflection on the diversity of recommended content [12].

### 4.2 RQ2: User Modeling Methods

The ethical concerns surrounding UGC recommendation and moderation algorithms underscore the importance of accurately modeling users. In this section, we summarize HCI research that approaches user modeling from user-centered perspectives. Table 2 presents the six user modeling subthemes and their definitions, and Table 5 lists the included papers. Figure 4-left, provides an overview of the distribution, and Figure 4-right, illustrates the number of papers referencing each combination of ethical concerns and user modeling methods. For this review, we focus on user modeling approaches that are discussed in at least three papers within each ethical dimension.

*4.2.1 User and Content Characterization.* HCI research has employed user and content characterization, involving the modeling of users through analysis of UGC or profile features, to provide insights into how RecSys reflects consumers' true demands. Thirty-nine papers have explored methods for user modeling by examining both explicit and implicit preferences, as well as behaviors associated with the creation and consumption of UGC. Several studies employ content-based analysis of multi-modal UGC, including text, images, and videos to gain a deeper understanding of user preferences [5, 9, 26, 98]. Other research focus on profiling and modeling users by analyzing their behaviors, including content creation, commenting, preference settings, and interaction logs [2, 11, 69, 86]. Techniques such as topic modeling, collaborative filtering [9, 11, 41], and tensor-based contextual modeling [52] have been proposed to build user profiles and capture implicit interests through social contexts.

Eighteen papers employ data characterization methods to model users' *interest and engagement*. HCI research identifies three main categories of user behaviors for capturing user values. First, behavioral signals such as clicks, likes, views, and comments are commonly used to infer user interests [4, 5, 98, 114]. Second, modeling social and network-based interests helps reveal the values of user groups and communities [16, 27, 145]. Third, profile-based and content-derived models—such as bios [2], profile-content similarity [9], and past collaborative activities [41]—offer additional insight into user interests.

Nine papers specifically address recommendation *diversity* by characterizing UGC and user profiles. HCI research has examined diversification criteria based on UGC content [4, 59, 98] and user demographics [88] to enhance recommendation diversity. These studies explore techniques to filter redundant content, mitigate biases, balance representation [66, 80, 88], recommend diverse social connections [118], and improve geographical, contextual, and real-time diversity [58, 59, 114].

Five papers focus on improving *transparency* in user modeling through explicitly quanifying the biases in user modeling [43, 80, 86]. Additionally, various studies have proposed user interface designs to present user modeling data and algorithmic explanations in more accessible and user-friendly formats [16, 24, 46].

Thirteen papers investigate methods for characterizing UGC content and user behaviors to improve content *Adversarial Content Dynamics* on UGC platforms. These studies analyze fake or manipulative accounts [17, 127], misinformation [66, 70], and manipulative behaviors [85, 123, 136]. HCI research has also focused on characterizing harmful content [26, 81] and behaviors aimed at circumventing moderation [53].

Six papers specifically address *fairness* ethics by characterizing biases in UGCAlogs, output representativeness [43, 59, 80, 88], and audience targeting [15]. These studies emphasize the importance of fairness-aware algorithms that ensure equitable treatment of diverse user groups and content.

*4.2.2 Contextual Impact on Consumers.* HCI researchers have investigated how user experiences with UGC—such as intentions, emotions, and feelings—may vary across different contexts. Six papers investigate the impact of algorithm-driven content on user experiences using human-centered methods such as surveys, interviews, and design probes. These methods are used to examine UGCAlgos' influence on social connections with friends [111], their role in shaping trust in platforms [130], and their impact on learning efficiency when interacting with UGC [125]. Among studies using this modeling technique, three specifically investigate the social and emotional effects of UGCAlgos on user *interest and engagement* with content [12, 93, 133].

*4.2.3 Creators' and Consumers' Folk Thoery.* In HCI, as examined in 22 papers, researchers have categorized how content creators and consumers develop different informal beliefs (i.e., *folk theories*) about how UGCAlgos function. HCI research investigates the formation of algorithmic folk theories [13] and their influence on content consumers [77, 111] through both large-scale computational methods and in-depth human subject research. These approaches further reveal additional concerns include human-algorithm collaboration in moderation practices [94, 95, 129] and both human and algorithm biases against marginalized creators [74].

Eight papers explore users' perceptions of algorithm *transparency*. Creators and consumers often perceive UGCAlgos as "black boxes" that lack clear explanations, leading to uncertainty and distrust regarding how algorithms influence their content or experiences [44, 46, 94, 111, 129]. To navigate the opacity of UGCAlgos, users often develop informal explanations, associating algorithms with anecdotal behavior patterns and even attributing personalities to them [44, 139]. Studies suggest that users express a strong desire for greater transparency to enable more effective interactions with these algorithms [44, 111, 129]. In response, research has proposed various frameworks and mechanisms to improve users' perceptions of algorithmic transparency [30, 112].

Four papers discuss diverse users' beliefs about content *Adversarial Content Dynamics* regarding UGC recommendation and moderation. Users are found to form beliefs about content appropriateness by evaluating whether the recommended UGC aligns with their personal values [93, 139] or whether inappropriate content has been effectively moderated [30, 56].

Seventeen papers examine users' beliefs regarding the *fairness* of UGCAlgos. In those studies, most users perceive a lack of transparency in UGCAlgos, leading to outcomes they view as unfair or unpredictable [44, 94, 95, 129]. HCI studies find that many users believe UGCAlgos exhibit biases against certain groups, potentially suppressing creators with specific identities [31, 39, 73, 74, 116, 122]. Concerns are also raised about unfair moderation policies [74, 95] and a perceived lack of user control over UGCAlgos [19, 44, 116, 129]. Notably, providing clear explanations to creators can significantly enhance the perceived acceptability of algorithmic moderation [44, 56]. Additionally, users' beliefs about algorithmic fairness can also be influenced by their political ideologies and affiliations [19, 22].

*4.2.4 Creator Algorithm and Policy Outsmart.* There has been growing attention in HCI research to examining how content creators actively counteract recommendation or moderation outcomes in order to achieve their visibility and monetary goals. HCI research has revealed that creators sometimes strategically interact with UGC content and platforms based on their understanding of UGCAlgos, attempting to exert control over the algorithms despite the lack of explicit control features provided by the platforms. The phenomenon of users attempting to "outsmart" algorithms is explored in 14 papers, which highlight various tactics employed to bypass moderation [19, 51] and to circumvent algorithmic gatekeeping [82, 139].

Four papers discuss user strategies to outsmart algorithms that could harm the *Adversarial Content Dynamics* ethic. HCI research indicates that strategies designed to circumvent moderation can enable inappropriate content to bypass standard filters, often by exploiting inflated popularity metrics [51, 53, 136]. Furthermore, to avoid demonetization by algorithms, content creators often tailor their content to align with community standards and steer clear of material likely to be flagged as inappropriate [139].

Ten papers examine how users outsmart UGCAlgos in pursuit of *fairness* on platforms. UGC creators often adapt their content strategically to align with or manipulate algorithmic preferences [31, 94, 122]. Some creators disengage from certain platforms or communities as a response to perceived algorithmic unfairness [116]. Others form supportive communities to share strategies and ideas for coping with unfair algorithmic treatment [13, 33, 39, 73, 122].

*4.2.5 Recommendation Visualization and Control.* In our review, 14 papers propose new interface designs for UGCAlgo systems, focusing on the use of visualization and control mechanisms to provide consumers with novel recommendation experiences. Through user-centered design and evaluation process, HCI research plays an important role in eliciting latent user needs and requirements for UGCAlgos. A predominant focus is on visualization techniques for displaying recommended items [50, 59, 125, 126] and mapping user relationships [58].

Six papers have explored how interface design can enhance user *interest and engagement*. Visualization techniques have been employed to enable users to better refine their preferences [59, 125]. Some UGCAlgos allow users to provide feedback or adjust preferences directly [16, 23]. Additionally, interfaces can incorporate social interactions to facilitate the discovery of new interests [12]. For creators, UGCAlgo interface design can improve their engagement by giving them greater control over who sees their content [46].

Six papers explore interface designs for encouraging the exploration of *diversity* in recommendations. Interfaces are employed to present diverse opinions, content from different geographical locations [50, 59], and other users [12]. Interactive tools that incorporate visualization techniques have been developed to present a wide array of content [58, 126].

Five papers have explored ways to improve UGCAlgo *transparency* through interface design. The proposed interfaces aim to make recommendation algorithms more explicit and understandable for UGC users [3, 16, 44]. For content creators, specific interfaces have been designed to support understanding of content delivery processes [46] and moderation decisions [86].

Three papers have examined interface designs that address the *fairness* of UGCAlgos. These HCI studies propose interfaces designed to present fairly selected content and balanced opinions [44, 50, 59].

*4.2.6 Community and Societal Impact.* Our review includes 19 articles that focus on specific communities and examine the broader community and societal impacts of UGCAlgos beyond the individual level. These studies investigate how algorithmic processes can perpetuate and amplify existing social inequities, leading to the marginalization of certain groups [73, 74, 94, 116].

Four papers explore *diversity* in the context of community and societal dynamics. These studies specifically argue that UGCAlgos can overrepresent content from central locations [59], contribute to the formation of echo chambers and filter bubbles on social media [42, 51, 60], and mobilize like-minded individuals against marginalized groups [51].

Five papers examine the impact of UGCAlgo's delivery of *inappropriate content* on communities and society. HCI research has investigated sexual risks, harassment, and abuse in recommended UGC that threatens underprivileged youth [20] and women [97]. Studies have also highlighted hate speech and discriminatory language targeting immigrants [45], as well as religious and political minority groups [35, 51].

Fourteen papers explore community and social dynamics related to the *fairness* of UGCAlgos. These studies examine how UGCAlgos can hinder creators with disabilities from achieving full and effective participation in society [31, 116]. For creators in LGBTQ+ groups, HCI research has investigated their perceptions of identity [73], challenges navigating unfair visibility [39, 122], and experiences with demonetization [74]. Other groups experiencing algorithmic unfairness include women [54], decentralized groups [35, 59], and immigrants [45].

## 5 Discussion

While much of the research on conventional RecSys has focused on addressing ethical challenges during data collection, algorithm design, and deployment [37, 72, 99, 120], our review highlights the unique ethical aspects of the UGC context explored in HCI research, emphasizing how human-centered approaches can model users' expectations of ethical algorithms. This section summarizes six key themes through which ethical considerations are uniquely defined and studied in HCI research.

### 5.1 Mitigating Adversarial UGC Through Understanding Human Behavior

Milano's framework [100] defines *inappropriate content* in RecSys as content that conflicts with users' values, highlighting the importance of filtering techniques. RecSys research also identifies key concerns such as misinformation and toxic language targeting specific groups [36, 137]. Although other reviews have noted that users can manipulate RecSys [120], assigning accountability for harmful content remains a significant challenge [37, 72].

Our review of HCI literature suggests that concerns about the appropriateness of UGC primarily center on misleading content, hate speech and harassment, and health-harming material. However, evaluating the role of UGCAlgos must account for three dynamic human factors. First, because UGCAlgos are designed to maximize user attention and enable personalization, these features can amplify the dissemination of inappropriate content [66]. Personalization capabilities may connect like-minded and maliciously motivated creators with consumers, facilitating the formation of funnels and communities that promote toxicity on UGC platforms [51, 81, 97, 106]. Second, while accountability is a central topic in RecSys [37, 72], HCI studies emphasize the importance of attending more closely to specific creator behaviors, such as the use of fake accounts [17], engagement pods [136], and clickbait tactics [85]. New forms of adversarial content continually emerge, requiring constant attention to and understanding of new online communities [26, 97, 106]. Third, efforts to bypass moderation are also dynamic, suggesting an ongoing attack-defense cycle between creators and moderation systems [56]. Platforms should therefore avoid relying solely on algorithmic moderation and instead integrate human moderators to strengthen these efforts [20].

To enhance content appropriateness, HCI research proposes several user modeling approaches. *User and Content Characterization* is essential for identifying misinformation [66] and tactics used to evade moderation [53]. This knowledge is critical for informing algorithmic design and policy development. Analyzing *Creators' and Consumers' Folk Thoery* about recommendation and moderation systems [93, 139] is key to balancing personal interests with potential harm. There is also a need for an evolving understanding of *Creator Algorithm and Policy Outsmart* strategies, as users continuously develop new methods to boost visibility or bypass moderation [97, 139]. Additionally, examining *Community and Societal Impact* responses to inappropriate content reveals group-specific needs for algorithmic safeguards [20, 97].

Future work on mitigating inappropriate UGC should address the critical balance between fostering user engagement and preventing the formation of toxic communities through UGCAlgos. One open direction is to enhance the harm-reduction functions of UGCAlgos so that they promote healthy engagement rather than toxicity or misinformation. Another challenge is that adversarial content will continually emerge, requiring the design of adaptive moderation approaches that meaningfully combine human and algorithmic efforts. At the same time, such moderation must ensure *Socio-Algorithmic Fairness*. The design of moderation in UGCAlgos should safeguard vulnerable users while avoiding the unintentional suppression of their legitimate expression.

### 5.2 Protecting User Privacy in Interface Design

*Privacy* in Milano's framework refers to technical safeguards that prevent data leakage or privacy violations. Prior research on RecSys ethics has emphasized that incorporating users' behavioral and personal data carries significant risks for privacy breaches [38, 120]. The RecSys ethics literature has underscored the importance of regulating data misuse, overcollection, and unauthorized exposure as key strategies for addressing these concerns [72].

Although the HCI literature in our review pool does not extensively examine privacy concerns, a few system designs incorporate user-configurable settings that allow users to control how UGCAlgos utilize personal information [46, 86, 119]. While our review shows that much of HCI research has focused on enhancing user engagement and detecting inappropriate content, the implications of these practices for user privacy remain underexplored and warrant further investigation.

A pressing need is to examine the balance between content creators' desire for engagement and their concerns about overexposure on UGC platforms. New human-centered mechanisms should be designed to address privacy control needs, enabling users to carefully manage the visibility of their content. Another future direction, as suggested in prior work [46, 86], is to integrate privacy considerations into the *Transparency Structures*, making the use of personal information more transparent to both content creators and consumers.

### 5.3 Capturing True User Interest and Enhancing Healthy Engagement

*Autonomy and Personal Identity* is framed by Milano as maintaining user control and avoiding addictive behaviors [100]. Prior research on RecSys systems has examined how recommendations, beyond driving short-term engagement, can lead to unintended consequences such as addictive behaviors [99, 120]. Therefore, preserving user autonomy is recognized as essential [38, 72].

HCI studies emphasize how UGCAlgo designs can better reflect users' authentic selves and values, rather than merely maximizing engagement or fostering addictive behaviors [46, 84, 101]. Human-centered techniques leverage implicit signals from natural interactions with UGC, such as clicks [119, 143] and browsing history [2, 29, 64, 114, 148, 149], to model user preferences. Social network information has also been used to enhance personalization and interaction [9, 11, 41, 90, 134, 142, 145]. UGCAlgos, however, face the challenge of balancing user interests with *Diversification*. While researchers strive to better capture preferences and boost engagement [46, 64, 101, 143, 149], overly personalized systems risk creating "filter bubbles" that reinforce biased opinions and limit exposure to diverse perspectives [42, 51, 88]. Consequently, HCI research stresses the need to capture users' underlying values across posting, consumption, and friending behaviors [46, 84, 101] and to ensure fair representation of diverse interests while mitigating filter bubble effects [25, 98].

User modeling methods for promoting healthy engagement with UGC include *User and Content Characterization* techniques such as leveraging multimodal user data [4, 98], social network data [27, 145], and profile data [2] to develop a comprehensive interpretation of user interests. HCI research highlights that user context—including intentions, emotions, and content discovery—is a critical factor in understanding what users truly value [12, 93]. In addition, *Recommendation Visualization and Control* such as data visualizations has the potential to encourage users to more effectively explore and configure their interests [58, 125], and to better align UGCAlgos with the dynamic and evolving needs of users [23].

Moving forward, one open challenge for future HCI and RecSys research is to develop methods for personalization with diversity, ensuring recommendations promote discovery and exposure to new perspectives rather than reinforcing filter bubbles. This requires addressing the potential short-term engagement loss when introducing diverse content. One approach is to design interfaces that support reflection, enabling users to adjust their own consumption patterns [12]. Another direction is to integrate human signals such as emotion, intent, and social dynamics into RecSys while building the *Transparency Structure* to enhance user understanding of these factors. Addressing these challenges requires advancing RecSys modeling in tandem with HCI innovations.

### 5.4 Providing Transparency Beyond Explaining Recommendation Algorithms

Milano's framework highlights the issue of opacity in RecSys and the need to enhance transparency by providing explanations [100]. Similarly, RecSys research often discusses transparency and explainability in terms of the "black box" nature of recommendation logic, which can undermine user trust and accountability [37, 38, 72, 99].

In HCI research, while explanatory transparency of UGCAlgos remains central, scholars have called for broader transparency in the UGC ecosystem—including contestability and control transparency—which is critical not only for content consumers but also for creators and platform governance. Contestability transparency enables creators to challenge algorithmic decisions through platform mechanisms and supports their pursuit of fair treatment [129, 130]. Control transparency allows users to understand how their actions influence algorithmic outputs [112]. This finding suggests that HCI research calls for UGC platforms to establish structured transparency. A holistic transparency framework can help prevent the development of inaccurate folk theories about algorithmic behavior [44, 139] and is essential for promoting fairness and trust in visibility-related matters [31, 43].

A human-centered user modeling approach to addressing UGCAlgo transparency involves examining *Creators' and Consumers' Folk Thoeries* about these algorithms to bridge the gap between user expectations and system operations [44, 139]. HCI studies have also proposed methods to characterize potential biases in user and content [43, 80]. Within *Recommendation Visualization and Control*, researchers have developed approaches to help both creators and users better understand recommendation and moderation outcomes [16, 86].

Future research on transparency should move beyond explanation toward richer forms of contestability and control. An open question is how to design mechanisms that allow creators to meaningfully contest or influence algorithmic decisions without overwhelming them or degrading system performance. Another challenge is balancing trust and workload: while increased transparency can empower users, excessive explanations or preference prompts may erode usability and engagement. One potential direction is to integrate *Creators' and Consumers' Folk Theories* into the design of transparency tools, enabling UGC platforms to build complementary transparency structures that strengthen UGC RecSys.

### 5.5 Ensuring Fairness in Visibility, Moderation, and User Interaction

*Fairness* in Milano's framework [100] emphasizes the avoidance of algorithmic bias in data ranking. RecSys research on fairness has primarily focused on ensuring equitable treatment across users and items [146]. For example, RecSys must provide users with equal access to relevant content [37, 72] and balance the needs and rights of diverse stakeholders [37, 99]. Biased recommendations can reinforce existing social biases and discrimination [120].

In our review, HCI researchers also emphasized the socio-technical aspects of fairness in recommendations—specifically, ensuring that creators from diverse groups are fairly represented. For example, allocative fairness, which is often compromised not only by recommendation algorithms but also by moderation algorithms, significantly affects creators' visibility and their willingness to contribute creative labor [13, 18, 95]. Creators have called for fair governance that includes protections for free expression, avoidance of excessive censorship [56, 95], consistency in moderation practices [35], and a fair appeals process [129, 130]. HCI research has also highlighted that when user interactions are not accurately reflected in algorithmic outputs, it undermines users' perceptions of fairness in UGCAlgos [44, 111].

User modeling methods for addressing fairness focus on understand the beliefs of diverse user groups. HCI studies show that *Creators' and Consumers' Folk Thoery* such as predictability [44], policy transparency and explanation [44, 56, 74, 95], and political stance [19, 22] shape users' perceptions of fairness in UGCAlgos. HCI research also shows that fairness impacts identity formation and monetization on UGC platforms [73, 74]. *Creator Algorithm and Policy Outsmart* activities illustrate how fairness influences creative labor [31, 95] and encourage creators to share tactics in pursuit of fair treatment [33, 122]. To support fairness, designers require *User and Content Characterization* methods to evaluate the severity of biases in UGCAlgos [59, 80].

Fairness in UGC platforms requires bridging RecSys advances with HCI's focus on social–algorithmic fairness. An important open direction is reconciling allocative fairness while minimizing risks of amplifying *Adversarial Content Dynamics*. Another challenge is designing governance mechanisms—including moderation practices, appeals processes, and transparency tools—that balance stakeholder rights yet remain scalable within algorithmic systems. Future work should pursue adaptive fairness frameworks that integrate RecSys methods for bias detection with HCI insights on *Transparency Structure* and *Diversification*. Transparency in allocation and moderation is an integral part of platform governance, while diversification can broaden the promotion of creators' content by UGCAlgos.

### 5.6 Incorporating Diversity Evaluation Metrics and Social-Based and Interactive Diversification

Milano's framework highlights that RecSys have transformative *social effects* because they can reinforce existing biases and contribute to the formation of filter bubbles, thereby underscoring the need to enhance the diversity of recommendations [100]. Methodologically, algorithm-centric RecSys research has primarily focused on reducing intra-list similarity and promoting serendipity [146]. Other studies addressing this ethical dimension emphasize how RecSys algorithms can expose users to novel and diverse content to avoid over-optimization and excessive personalization [38, 120].

In contrast, HCI researchers have not only proposed improving algorithms to enhance content diversity but have also raised questions about how to incorporate various social and interactive methods to support the exploration of diverse content. Algorithmic diversification focuses on auditing the diversity levels of UGCAlgos [8, 66] and developing diversity evaluation metrics [80, 104, 109]. HCI researchers have also noted that social factors, such as upvoting mechanisms [51] and users' geolocation [58], can influence the diversity of UGCAlgos and help reduce filter bubbles [42, 51, 88]. In designs supporting interactive exploration, HCI studies find that users can enhance their engagement with diverse content by reflecting on the relationships between recommended items [50] or by interacting with recommendations from different individuals [12].

Human-centered methods include *User and Content Characterization* approaches that develop criteria for content diversity [4, 59] and measure the biases, social, and contextual factors that hinder diversity [59, 80, 118]. From an interaction perspective, HCI researchers have noted that *Recommendation Visualization and Control* can be designed to present diverse opinions, with an emphasis on promoting reflection on content diversity [16, 148] and encouraging discovery [12, 126]. Additionally, in studies examining this dimension within *Community and Societal Impact*, HCI research highlights that a lack of diversity in UGCAlgos can contribute to the formation of echo chambers and the mobilization of radical groups [42, 51].

For future research, a central open question is how to balance algorithmic diversification with user agency. Researchers should examine how UGCAlgos can either passively inject diverse content or actively invite users to explore it through interactive interfaces. Another challenge is understanding the *Contextual Impact on Users*. Diversification should not be based solely on statistical variety but also on social value, cultural representation, and community needs. Such approaches could also potentially help mitigate filter bubbles and polarization.

## 6 Conclusion and Future Work

Our descriptive review provides key insights into the ethical challenges associated with UGCAlgos and highlights human-centered perspectives for addressing these challenges. These concerns are more deeply intertwined in UGCAlgos due to the dual role of users as both content creators and consumers. This duality amplifies the complexity of ethical challenges, as the algorithms must simultaneously navigate the responsibilities of promoting fair content moderation and dissemination, ensuring equitable visibility, and safeguarding user experiences in content consumption.

Future research should explore ways to enhance user experiences, guide algorithmic development, and craft effective policies that align with user beliefs and needs. While this review illuminates the ethical dimensions of UGCAlgos, it primarily focuses on UGC platforms, leaving opportunities for broader exploration in diverse social media contexts such as Wikipedia, MOOCs, sharing economy platforms, and job-sharing sites. Furthermore, the review does not comprehensively cover all articles in machine learning and AI, potentially overlooking contributions related to human factors and user experience. As AI continues to play an expanding role in social media, future studies could apply our framework across various contexts and disciplines, further enriching our understanding of UGCAlgos and their societal implications.

## References

[1] Behnoush Abdollahi and Olfa Nasraoui. 2018. Transparency in fair machine learning: The case of explainable recommender systems. In *Proceedings of the Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent*, Jianlong Zhou and Fang Chen (Eds.). Springer International Publishing, Cham, 21–35. DOI: https://doi.org/10.1007/978-3-319-90403-0{_}2

[2] Al-Batool Al-Ghamdi, Ameenah Al-Sulami, Nouf Al-Jadani, and Maha Aljohani. 2020. Support vector machine algorithm to classify instagram users' accounts based on users' interests, Helmut Degen and Lauren Reinerman-Jones (Eds.). Springer International Publishing, Cham, 179–196.

[3] Oscar Alvarado and Annika Waern. 2018. Towards algorithmic experience: Initial efforts for social media contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18)*. Association for Computing Machinery, New York, NY, USA, 1–12. DOI: https://doi.org/10.1145/3173574.3173860

[4] Maryam Aziz, Jesse Anderton, Kevin Jamieson, Alice Wang, Hugues Bouchard, and Javed Aslam. 2022. Identifying new podcasts with high general appeal using a pure exploration infinitely-armed bandit strategy. In *Proceedings of the 16th ACM Conference on Recommender Systems (RecSys'22)*. Association for Computing Machinery, New York, NY, USA, 134–144. DOI: https://doi.org/10.1145/3523227.3546766

[5] Saeideh Bakhshi, David A Shamma, Lyndon Kennedy, Yale Song, Paloma de Juan, and Joseph 'Jofish' Kaye. 2016. Fast, cheap, and good: Why animated GIFs engage Us. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'16)*. Association for Computing Machinery, New York, NY, USA, 575–586. DOI: https://doi.org/10.1145/2858036.2858532

[6] Giacomo Balloccu, Ludovico Boratto, Gianni Fenu, and Mirko Marras. 2022. Post processing recommender systems with knowledge graphs for recency, popularity, and diversity of explanations. In *SIGIR 2022 - Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval* (2022), 646–656. DOI: https://doi.org/10.1145/3477495.3532041/SUPPL{_}FILE/

[7] Krisztian Balog, Filip Radlinski, and Shushan Arakelyan. 2019. Transparent, scrutable and explainable user models for personalized recommendation. In *SIGIR 2019 - Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (2019), 265–274. DOI: https://doi.org/10.1145/3331184.3331211/SUPPL{_}FILE/CITE1-16H40-D1.MP4

[8] Jack Bandy and Nicholas Diakopoulos. 2021. More accounts, fewer links: How algorithmic curation impacts media exposure in twitter timelines. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (4 2021). DOI: https://doi.org/10.1145/3449152

[9] Trapit Bansal, Mrinal Das, and Chiranjib Bhattacharyya. 2015. Content driven user profiling for comment-worthy recommendations of news and blog articles. In *Proceedings of the 9th ACM Conference on Recommender Systems (RecSys'15)*. Association for Computing Machinery, New York, NY, USA, 195–202. DOI: https://doi.org/10.1145/2792838.2800186

[10] Massimo Bartoletti, Stefano Lande, and Alessandro Massa. 2016. Faderank: An incremental algorithm for ranking twitter users, Wojciech Cellary, Mohamed F Mokbel, Jianmin Wang, Hua Wang, Rui Zhou, and Yanchun Zhang (Eds.). Springer International Publishing, Cham, 55–69.

[11] Preeti Bhargava, Oliver Brdiczka, and Michael Roberts. 2015. Unsupervised modeling of users' interests from their facebook profiles and activities. In *Proceedings of the 20th International Conference on Intelligent User Interfaces (IUI'15)*. Association for Computing Machinery, New York, NY, USA, 191–201. DOI: https://doi.org/10.1145/2678025.2701365

[12] Md Momen Bhuiyan, Carlos Augusto Bautista Isaza, Tanushree Mitra, and Sang Won Lee. 2022. OtherTube: Facilitating content discovery and reflection by exchanging youtube recommendations with strangers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI'22)*. Association for Computing Machinery, New York, NY, USA. DOI: https://doi.org/10.1145/3491102.3502028

[13] Sophie Bishop. 2019. Managing visibility on YouTube through algorithmic gossip. *New Media & Society* 21, 11-12 (6 2019), 2589–2606. DOI: https://doi.org/10.1177/1461444819854731

[14] J Bobadilla, F Ortega, A Hernando, and A Gutiérrez. 2013. Recommender systems survey. *Knowledge-Based Systems* 46 (2013), 109–132. DOI: https://doi.org/10.1016/j.knosys.2013.03.012

[15] Nadine Bol, Joanna Strycharz, Natali Helberger, Bob van de Velde, and Claes H de Vreese. 2020. Vulnerability in a tracked society: Combining tracking and survey data to understand who gets targeted with what content. *New Media & Society* 22, 11 (10 2020), 1996–2017. DOI: https://doi.org/10.1177/1461444820924631

[16] Svetlin Bostandjiev, John O'Donovan, and Tobias Höllerer. 2012. TasteWeights: A visual interactive hybrid recommender system. In *Proceedings of the 6th ACM Conference on Recommender Systems (RecSys'12)*. Association for Computing Machinery, New York, NY, USA, 35–42. DOI: https://doi.org/10.1145/2365952.2365964

[17] Adam Breuer, Roee Eilat, and Udi Weinsberg. 2020. Friend or faux: Graph-based early detection of fake accounts on social networks. In *Proceedings of the Web Conference 2020 (WWW'20)*. Association for Computing Machinery, New York, NY, USA, 1287–1297. DOI: https://doi.org/10.1145/3366423.3380204

[18] Taina Bucher. 2012. Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society* 14, 7 (4 2012), 1164–1180. DOI: https://doi.org/10.1177/1461444812440159

[19] Jenna Burrell, Zoe Kahn, Anne Jonas, and Daniel Griffin. 2019. When users control the algorithms: Values expressed in practices on twitter. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019). DOI: https://doi.org/10.1145/3359240

[20] Xavier V Caddle, Nurun Naher, Zachary P Miller, Karla Badillo-Urquiola, and Pamela J Wisniewski. 2022. Duty to respond: The challenges social service providers face when charged with keeping youth safe online. *Proceedings of the ACM on Human-Computer Interaction* 7, GROUP (12 2022). DOI: https://doi.org/10.1145/3567556

[21] André Calero Valdez, Martina Ziefle, and Katrien Verbert. 2016. HCI for recommender systems: The past, the present and the future. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys'16)*. Association for Computing Machinery, New York, NY, USA, 123–126. DOI: https://doi.org/10.1145/2959100.2959158

[22] Mikhaila N Calice, Luye Bao, Isabelle Freiling, Emily Howell, Michael A Xenos, Shiyu Yang, Dominique Brossard, Todd P Newman, and Dietram A Scheufele. 2021. Polarized platforms? How partisanship shapes perceptions of "algorithmic news bias". *New Media & Society* (8 2021), 14614448211034159. DOI: https://doi.org/10.1177/14614448211034159

[23] Amy Campbell, Christopher Wienberg, and Andrew Gordon. 2012. Collecting relevance feedback on titles and photographs in weblog posts. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces (IUI'12)*. Association for Computing Machinery, New York, NY, USA, 139–148. DOI: https://doi.org/10.1145/2166966.2166993

[24] Samara Castillo, Héctor Allende-Cid, Wenceslao Palma, Rodrigo Alfaro, Heitor S Ramos, Cristian Gonzalez, Claudio Elortegui, and Pedro Santander. 2019. Detection of bots and cyborgs in twitter: A study on the chilean presidential election in 2017, Gabriele Meiselwitz (Ed.). Springer International Publishing, Cham, 311–323.

[25] Abhijnan Chakraborty, Gourab K Patro, Niloy Ganguly, Krishna P Gummadi, and Patrick Loiseau. 2019. Equality of voice: Towards fair representation in crowdsourced top-k recommendations. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*'19)*. Association for Computing Machinery, New York, NY, USA, 129–138. DOI: https://doi.org/10.1145/3287560.3287570

[26] Stevie Chancellor, Yannis Kalantidis, Jessica A Pater, Munmun De Choudhury, and David A Shamma. 2017. Multimodal classification of moderated online pro-eating disorder content. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI'17)*. Association for Computing Machinery, New York, NY, USA, 3213–3226. DOI: https://doi.org/10.1145/3025453.3025985

[27] Shuo Chang, Vikas Kumar, Eric Gilbert, and Loren G Terveen. 2014. Specialization, homophily, and gender in a social curation site: Findings from pinterest. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work \& Social Computing (CSCW'14)*. Association for Computing Machinery, New York, NY, USA, 674–686. DOI: https://doi.org/10.1145/2531602.2531660

[28] Lijiang Chen, Yibing Zhao, Shimin Chen, Hui Fang, Chengkai Li, and Min Wang. 2013. iPLUG: Personalized list recommendation in twitter, Xuemin Lin, Yannis Manolopoulos, Divesh Srivastava, and Guangyan Huang (Eds.). Springer Berlin , Berlin, 88–103.

[29] Minmin Chen, Can Xu, Vince Gatto, Devanshu Jain, Aviral Kumar, and Ed Chi. 2022. Off-policy actor-critic for recommender systems. In *Proceedings of the 16th ACM Conference on Recommender Systems (RecSys'22)*. Association for Computing Machinery, New York, NY, USA, 338–349. DOI: https://doi.org/10.1145/3523227.3546758

[30] Shih-Yi Chien, Cheng-Jun Yang, and Fang Yu. 2022. XFlag: Explainable fake news detection model on social media. *International Journal of Human-Computer Interaction* 38, 18-20 (12 2022), 1808–1827. DOI: https://doi.org/10.1080/10447318.2022.2062113

[31] Dasom Choi, Uichin Lee, and Hwajung Hong. 2022. "It's not wrong, but I'm quite disappointed": Toward an inclusive algorithmic experience for content creators with disabilities. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[32] W George Kernohan3 PhD CPhys CMath. 2011. Ethical issues in undertaking internet research of user-generated content: A review of the literature. *Evidence Based Midwifery* (2011).

[33] Kelley Cotter. 2018. Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society* 21, 4 (12 2018), 895–913. DOI: https://doi.org/10.1177/1461444818815684

[34] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for YouTube recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys'16)*. Association for Computing Machinery, New York, NY, USA, 191–198. DOI: https://doi.org/10.1145/2959100.2959190

[35] Dipto Das, Carsten \Osterlund, and Bryan Semaan. 2021. "Jol" or "pani"?: How does governance shape a platform's identity? *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (10 2021). DOI: https://doi.org/10.1145/3479860

[36] Yashar Deldjoo, Zhankui He, Julian McAuley, Anton Korikov, Scott Sanner, Arnau Ramisa, Rene Vidal, Maheswaran Sathiamoorthy, Atoosa Kasrizadeh, and Silvia Milano. 2024. Recommendation with generative models. arXiv:2409.15173. Retrieved from https://arxiv.org/abs/2409.15173 (2024).

[37] Yashar Deldjoo, Dietmar Jannach, Alejandro Bellogin, Alessandro Difonzo, and Dario Zanzonelli. 2024. Fairness in recommender systems: Research landscape and future directions. *User Modeling and User-Adapted Interaction* 34, 1 (2024), 59–108. DOI: https://doi.org/10.1007/s11257-023-09364-z

[38] Yashar Deldjoo, Markus Schedl, and Peter Knees. 2024. Content-driven music recommendation: Evolution, state of the art, and challenges. *Computer Science Review* 51 (2024), 100618. DOI: https://doi.org/10.1016/j.cosrev.2024.100618

[39] Michael Ann DeVito. 2022. How transfeminine TikTok creators navigate the algorithmic trap of visibility via folk theorization. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (11 2022). DOI: https://doi.org/10.1145/3555105

[40] Michael A DeVito, Jeremy Birnholtz, Jeffery T Hancock, Megan French, and Sunny Liu. 2018. How people form folk theories of social media feeds and what it means for how we study self-presentation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18)*. Association for Computing Machinery, New York, NY, USA, 1–12. DOI: https://doi.org/10.1145/3173574.3173694

[41] Pierpaola Di Bitonto, Maria Laterza, Teresa Roselli, and Veronica Rossano. 2013. Recommendation of collaborative activities in e-learning environments, Masaaki Kurosu (Ed.). Springer Berlin, Berlin, 484–492.

[42] Tim Donkers and Jürgen Ziegler. 2021. The dual echo chamber: Modeling social media polarization for interventional recommending. In *Proceedings of the 15th ACM Conference on Recommender Systems (RecSys'21)*. Association for Computing Machinery, New York, NY, USA, 12–22. DOI: https://doi.org/10.1145/3460231.3474261

[43] Arun Dunna, Katherine A Keith, Ethan Zuckerman, Narseo Vallina-Rodriguez, Brendan O'Connor, and Rishab Nithyanand. 2022. Paying attention to the algorithm behind the curtain: Bringing transparency to YouTube's demonetization algorithms. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (11 2022). DOI: https://doi.org/10.1145/3555209

[44] Malin Eiband, Sarah Theres Völkel, Daniel Buschek, Sophia Cook, and Heinrich Hussmann. 2019. When people and algorithms meet: User-reported problems in intelligent everyday applications. In *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI'19)*. Association for Computing Machinery, New York, NY, USA, 96–106. DOI: https://doi.org/10.1145/3301275.3302262

[45] Claire Stravato Emes and Arul Chib. 2022. Co-opted marginality in a controlled media environment: The influence of social media affordances on the immigration discourse. *ACM Transactions on Social Computing* 5, 1–4 (11 2022). DOI: https://doi.org/10.1145/3532103

[46] Sindhu Kiranmai Ernala, Stephanie S Yang, Yuxi Wu, Rachel Chen, Kristen Wells, and Sauvik Das. 2021. Exploring the utility versus intrusiveness of dynamic audience selection on facebook. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (10 2021). DOI: https://doi.org/10.1145/3476083

[47] Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I "like" It, then I hide it: Folk theories of social feeds. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'16)*. Association for Computing Machinery, New York, NY, USA, 2371–2382. DOI: https://doi.org/10.1145/2858036.2858494

[48] W van der V Soude Fazeli, Hendrik Drachsler, Marlies Bitter-Rijpkema, Francis Brouns, and P B Sloep. 2016. Accuracy is just the tip of the iceberg: A data-centric vs. user-centric evaluation. *IEEE Transactions on Learning Technologies* (2016).

[49] Lu Gan, Diana Nurbakova, Léa Laporte, and Sylvie Calabretto. 2020. Enhancing recommendation diversity using determinantal point processes on knowledge graphs. *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* 20 (2020), 2001–2004. DOI: https://doi.org/10.1145/3397271.3401213

[50] Mingkun Gao, Hyo Jin Do, and Wai-Tat Fu. 2018. Burst your bubble! an intelligent system for improving awareness of diverse social opinions. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI'18)*. Association for Computing Machinery, New York, NY, USA, 371–383. DOI: https://doi.org/10.1145/3172944.3172970

[51] Tiana Gaudette, Ryan Scrivens, Garth Davies, and Richard Frank. 2020. Upvoting extremism: Collective identity formation and the extreme right on reddit. *New Media & Society* 23, 12 (9 2020), 3491–3508. DOI: https://doi.org/10.1177/1461444820958123

[52] Hancheng Ge, James Caverlee, and Haokai Lu. 2016. TAPER: A contextual tensor-based approach for personalized expert recommendation. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys'16)*. Association for Computing Machinery, New York, NY, USA, 261–268. DOI: https://doi.org/10.1145/2959100.2959151

[53] Ysabel Gerrard. 2018. Beyond the hashtag: Circumventing content moderation on social media. *New Media & Society* 20, 12 (5 2018), 4492–4511. DOI: https://doi.org/10.1177/1461444818776611

[54] Ysabel Gerrard and Helen Thornham. 2020. Content moderation: Social media's sexist assemblages. *New Media & Society* 22, 7 (7 2020), 1266–1286. DOI: https://doi.org/10.1177/1461444820912540

[55] Emilia Gómez, Vicky Charisi, and Stephane Chaudron. 2021. Evaluating recommender systems with and for children: Towards a multi-perspective framework. (2021). Retrieved from https://www.statista.com/statistics/1150571/share-us-parents-young-child-watch-youtube-videos/

[56] João Gonçalves, Ina Weber, Gina M Masullo, Marisa Torres da Silva, and Joep Hofhuis. 2021. Common sense or censorship: How algorithmic moderators and message type influence perceptions of online content deletion. *New Media & Society* (7 2021), 14614448211032310. DOI: https://doi.org/10.1177/14614448211032310

[57] Robert Gorwa, Reuben Binns, and Christian Katzenbach. 2020. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society* 7, 1 (1 2020), 2053951719897945. DOI: https://doi.org/10.1177/2053951719897945

[58] Eduardo Graells-Garrido, Mounia Lalmas, and Ricardo Baeza-Yates. 2016. Data portraits and intermediary topics: Encouraging exploration of politically diverse profiles. In *Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI'16)*. Association for Computing Machinery, New York, NY, USA, 228–240. DOI: https://doi.org/10.1145/2856767.2856776

[59] Eduardo Graells-Garrido, Mounia Lalmas, and Ricardo Baeza-Yates. 2016. Encouraging diversity- and representation-awareness in geographically centralized content. In *Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI'16)*. Association for Computing Machinery, New York, NY, USA, 7–18. DOI: https://doi.org/10.1145/2856767.2856775

[60] Quentin Grossetti, Cédric du Mouza, and Nicolas Travers. 2019. Community-based recommendations on twitter: Avoiding the filter bubble, Reynold Cheng, Nikos Mamoulis, Yizhou Sun, and Xin Huang (Eds.). Springer International Publishing, Cham, 212–227.

[61] C Guo, X Tian, and T Mei. 2014. User specific friend recommendation in social media community. In *Proceedings of the 2014 IEEE International Conference on Multimedia and Expo (ICME)*. 1–6. DOI: https://doi.org/10.1109/ICME.2014.6890217

[62] Ido Guy, Inbal Ronen, Elad Kravi, and Maya Barnea. 2016. Increasing activity in enterprise online communities using content recommendation. *ACM Transactions on Computer-Human Interaction* 23, 4 (8 2016). DOI: https://doi.org/10.1145/2910581

[63] Félix Hernández del Olmo and Elena Gaudioso. 2008. Evaluation of recommender systems: A new approach. *Expert Systems with Applications* 35, 3 (2008), 790–804. DOI: https://doi.org/10.1016/j.eswa.2007.07.047

[64] Cheng Hsu and Cheng-Te Li. 2021. RetaGNN: Relational temporal attentive graph neural networks for holistic sequential recommendation. In *Proceedings of the Web Conference 2021 (WWW'21)*. Association for Computing Machinery, New York, NY, USA, 2968–2979. DOI: https://doi.org/10.1145/3442381.3449957

[65] Shangrong Huang, Jian Zhang, Shiyang Lu, and Xian-Sheng Hua. 2015. Social friend recommendation based on network correlation and feature co-clustering. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (ICMR'15)*. Association for Computing Machinery, New York, NY, USA, 315–322. DOI: https://doi.org/10.1145/2671188.2749325

[66] Eslam Hussein, Prerna Juneja, and Tanushree Mitra. 2020. Measuring misinformation in video search platforms: An audit study on YouTube. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (5 2020). DOI: https://doi.org/10.1145/3392854

[67] F O Isinkaye, Y O Folajimi, and B A Ojokoh. 2015. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal* 16, 3 (2015), 261–273. DOI: https://doi.org/10.1016/j.eij.2015.06.005

[68] Benjamin N Jacobsen. 2021. Regimes of recognition on algorithmic media. *New Media & Society* (10 2021), 14614448211053555. DOI: https://doi.org/10.1177/14614448211053555

[69] Amin Javari, Zhankui He, Zijie Huang, Raj Jeetu, and Kevin Chen-Chuan Chang. 2020. Weakly supervised attention for hashtag recommendation using graph data. In *Proceedings of the Web Conference 2020 (WWW'20)*. Association for Computing Machinery, New York, NY, USA, 1038–1048. DOI: https://doi.org/10.1145/3366423.3380182

[70] Byungkyu Kang, John O'Donovan, and Tobias Höllerer. 2012. Modeling topic specific credibility on twitter. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces (IUI'12)*. Association for Computing Machinery, New York, NY, USA, 179–188. DOI: https://doi.org/10.1145/2166966.2166998

[71] Kawaljeet Kaur Kapoor, Kuttimani Tamilmani, Nripendra P. Rana, Pushp Patil, Yogesh K. Dwivedi, and Sridhar Nerur. 2018. Advances in social media research: Past, present and future. *Information Systems Frontiers* 20, 3 (2018), 531–558. DOI: https://doi.org/10.1007/s10796-017-9810-y

[72] E Karakolis, P F Oikonomidis, and D Askounis. 2022. Identifying and addressing ethical challenges in recommender systems. In *Proceedings of the 2022 13th International Conference on Information, Intelligence, Systems & Applications (IISA)*. 1–6. DOI: https://doi.org/10.1109/IISA56318.2022.9904386

[73] Nadia Karizat, Dan Delmonaco, Motahhare Eslami, and Nazanin Andalibi. 2021. Algorithmic folk theories and identity: How TikTok users co-produce knowledge of identity and engage in algorithmic resistance. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (10 2021). DOI: https://doi.org/10.1145/3476046

[74] Sara Kingsley, Proteeti Sinha, Clara Wang, Motahhare Eslami, and Jason I Hong. 2022. "Give everybody [..] a little bit more equity": Content creator perspectives and responses to the algorithmic demonetization of content associated with disadvantaged groups. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (11 2022). DOI: https://doi.org/10.1145/3555149

[75] Marios Koniaris, Giorgos Giannopoulos, Timos Sellis, and Yiannis Vasileiou. 2014. Diversifying microblog posts, Boualem Benatallah, Azer Bestavros, Yannis Manolopoulos, Athena Vakali, and Yanchun Zhang (Eds.). Springer International Publishing, Cham, 189–198.

[76] Joseph A Konstan and John Riedl. 2012. Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction* 22, 1 (2012), 101–123. DOI: https://doi.org/10.1007/s11257-011-9112-x

[77] Jennifer Krueckeberg. 2021. Youth and algorithmic memory: Co-producing personal memory on instagram, Matthias Rauterberg (Ed.). Springer International Publishing, Cham, 253–264.

[78] J Krumm, N Davies, and C Narayanaswami. 2008. User-generated content. *IEEE Pervasive Computing* 7, 4 (2008), 10–11. DOI: https://doi.org/10.1109/MPRV.2008.85

[79] Madhusree Kuanr and Puspanjali Mohapatra. 2021. Assessment methods for evaluation of recommender systems: A survey. *Foundations of Computing and Decision Sciences* 46, 4 (12 2021), 393–421. DOI: https://doi.org/10.2478/fcds-2021-0023

[80] Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P Gummadi, and Karrie Karahalios. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW'17)*. Association for Computing Machinery, New York, NY, USA, 417–432. DOI: https://doi.org/10.1145/2998181.2998321

[81] Ugur Kursuncu, Manas Gaur, Carlos Castillo, Amanuel Alambo, Krishnaprasad Thirunarayan, Valerie Shalin, Dilshod Achilov, I Budak Arpinar, and Amit Sheth. 2019. Modeling islamist extremist communications on social media using contextual dimensions: Religion, ideology, and hate. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019). DOI: https://doi.org/10.1145/3359253

[82] Alex Leavitt and John J Robinson. 2017. The role of information visibility in network gatekeeping: Information aggregation on reddit during crisis events. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW'17)*. Association for Computing Machinery, New York, NY, USA, 1246–1261. DOI: https://doi.org/10.1145/2998181.2998299

[83] Angela Y Lee, Hannah Mieczkowski, Nicole B Ellison, and Jeffrey T Hancock. 2022. The algorithmic crystal: Conceptualizing the self through algorithmic personalization on TikTok. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (11 2022). DOI: https://doi.org/10.1145/3555601

[84] Beibei Li, Beihong Jin, Xinzhou Dong, and Wei Zhuo. 2021. MULTIPLE: Multi-level user preference learning for list recommendation, Wenjie Zhang, Lei Zou, Zakaria Maamar, and Lu Chen (Eds.). Springer International Publishing, Cham, 221–236.

[85] Juliane A Lischka and Marcel Garz. 2021. Clickbait news and algorithmic curation: A game theory framework of the relation between journalism, users, and platforms. *New Media & Society* 25, 8 (7 2021), 2073–2094. DOI: https://doi.org/10.1177/14614448211027174

[86] Baoxi Liu, Peng Zhang, Yubo Shu, Zhengqing Guan, Tun Lu, Hansu Gu, and Ning Gu. 2022. Building a personalized model for social media textual content censorship. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (11 2022). DOI: https://doi.org/10.1145/3555657

[87] Fan Liu, Zhiyong Cheng, Huilin Chen, Yinwei Wei, Liqiang Nie, and Mohan Kankanhalli. 2022. Privacy-preserving synthetic data generation for recommendation systems. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'22)*. Association for Computing Machinery, New York, NY, USA, 1379–1389. DOI: https://doi.org/10.1145/3477495.3532044

[88] Ping Liu, Karthik Shivaram, Aron Culotta, Matthew A Shapiro, and Mustafa Bilgic. 2021. The interaction between political typology and filter bubbles in news recommendation algorithms. In *Proceedings of the Web Conference 2021 (WWW'21)*. Association for Computing Machinery, New York, NY, USA, 3791–3801. DOI: https://doi.org/10.1145/3442381.3450113

[89] S Liu and Z Chen. 2019. Sequential behavior modeling for next micro-video recommendation with collaborative transformer. In *Proceedings of the 2019 IEEE International Conference on Multimedia and Expo (ICME)*. 460–465. DOI: https://doi.org/10.1109/ICME.2019.00086

[90] Chun Lo, Emilie de Longueau, Ankan Saha, and Shaunak Chatterjee. 2020. Edge formation in social networks to nurture content creators. In *Proceedings of the Web Conference 2020 (WWW'20)*. Association for Computing Machinery, New York, NY, USA, 1999–2008. DOI: https://doi.org/10.1145/3366423.3380267

[91] Pasquale Lops, Marco de Gemmis, and Giovanni Semeraro. 2011. Content-based recommender systems: State of the art and trends. Springer US, Boston, MA, 73–105. DOI: https://doi.org/10.1007/978-0-387-85820-3{_}3

[92] Jie Lu, Dianshuang Wu, Mingsong Mao, Wei Wang, and Guangquan Zhang. 2015. Recommender system application developments: A survey. *Decision Support Systems* 74 (2015), 12–32. DOI: https://doi.org/10.1016/j.dss.2015.03.008

[93] Xing Lu, Zhicong Lu, and Changqing Liu. 2020. Exploring TikTok use and non-use practices and experiences in China. In *Proceedings of the International Conference on Human-computer Interaction*, Gabriele Meiselwitz (Ed.). Springer International Publishing, Cham, 57–70.

[94] Renkai Ma and Yubo Kou. 2021. "How advertiser-friendly is my video?": YouTuber's socioeconomic interactions with algorithmic content moderation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (10 2021). DOI: https://doi.org/10.1145/3479573

[95] Renkai Ma and Yubo Kou. 2022. "I'm not sure what difference is between their content and mine, other than the person itself": A study of fairness perception of content moderation on YouTube. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (11 2022). DOI: https://doi.org/10.1145/3555150

[96] Alex Martinez, Mihnea Tufis, and Ludovico Boratto. 2024. Unmasking privacy: A reproduction and evaluation study of obfuscation-based perturbation techniques for collaborative filtering. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'24)*. Association for Computing Machinery, New York, NY, USA, 1753–1762. DOI: https://doi.org/10.1145/3626772.3657858

[97] Adrienne Massanari. 2015. #Gamergate and the fappening: How reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society* 19, 3 (10 2015), 329–346. DOI: https://doi.org/10.1177/1461444815608807

[98] Philip J McParlane, Yashar Moshfeghi, and Joemon M Jose. 2014. "Nobody comes here anymore, it's too crowded"; predicting image popularity on flickr. In *Proceedings of the International Conference on Multimedia Retrieval (ICMR'14)*. Association for Computing Machinery, New York, NY, USA, 385–391. DOI: https://doi.org/10.1145/2578726.2578776

[99] Silvia Milano, Taddeo Mariarosaria, , and Luciano Floridi. 2021. Ethical aspects of multi-stakeholder recommendation systems. *The Information Society* 37, 1 (1 2021), 35–45. DOI: https://doi.org/10.1080/01972243.2020.1832636

[100] 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

[101] Smitha Milli, Luca Belli, and Moritz Hardt. 2021. From optimizing engagement to measuring value. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT'21)*. Association for Computing Machinery, New York, NY, USA, 714–722. DOI: https://doi.org/10.1145/3442188.3445933

[102] Jihad Mohammad, Farzana Quoquab, Ramayah Thurasamy, and Main Naser Alolayyan. 2020. The effect of user-generated content quality on brand engagement: The mediating role of functional and emotional values. *Journal of Electronic Commerce Research* 21, 1 (2020), 39–55.

[103] Teresa K Naab and Annika Sehl. 2016. Studies of user-generated content: A systematic review. *Journalism* 18, 10 (10 2016), 1256–1273. DOI: https://doi.org/10.1177/1464884916673557

[104] Matti Nelimarkka, Jean Philippe Rancy, Jennifer Grygiel, and Bryan Semaan. 2019. (Re)Design to mitigate political polarization: Reflecting habermas' ideal communication space in the united states of america and finland. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019). DOI: https://doi.org/10.1145/3359243

[105] Masashi Okubo and Shun Tamura. 2019. A proposal of video evaluation method using facial expression for video recommendation system BT - human interface and the management of information. information in intelligent systems, Sakae Yamamoto and Hirohiko Mori (Eds.). Springer International Publishing, Cham, 254–268.

[106] Kostantinos Papadamou, Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Michael Sirivianos. 2021. "How over is It?" Understanding the incel community on YouTube. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2 (10 2021). DOI: https://doi.org/10.1145/3479556

[107] Guy Paré and Spyros Kitsiou. 2017. Methods for literature reviews. In *Proceedings of the Handbook of eHealth Evaluation: An Evidence-based Approach [Internet]*. University of Victoria.

[108] B Patel, P Desai, and U Panchal. 2017. Methods of recommender system: A review. In *Proceedings of the 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*. 1–4. DOI: https://doi.org/10.1109/ICIIECS.2017.8275856

[109] Bibek Paudel and Abraham Bernstein. 2021. Random walks with erasure: Diversifying personalized recommendations on social and information networks. In *Proceedings of the Web Conference 2021 (WWW'21)*. Association for Computing Machinery, New York, NY, USA, 2046–2057. DOI: https://doi.org/10.1145/3442381.3449970

[110] Pearl Pu, Li Chen, and Rong Hu. 2011. A user-centric evaluation framework for recommender systems. In *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys'11)*. Association for Computing Machinery, New York, NY, USA, 157–164. DOI: https://doi.org/10.1145/2043932.2043962

[111] Emilee Rader. 2017. Examining user surprise as a symptom of algorithmic filtering. *International Journal of Human-Computer Studies* 98 (2017), 72–88. DOI: https://doi.org/10.1016/j.ijhcs.2016.10.005

[112] Emilee Rader, Kelley Cotter, and Janghee Cho. 2018. Explanations as mechanisms for supporting algorithmic transparency. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18)*. Association for Computing Machinery, New York, NY, USA, 1–13. DOI: https://doi.org/10.1145/3173574.3173677

[113] Guilherme Ramos, Mirko Marras, and Ludovico Boratto. 2024. Towards ethical item ranking: A paradigm shift from user-centric to item-centric approaches. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'24)*. Association for Computing Machinery, New York, NY, USA, 2667–2671. DOI: https://doi.org/10.1145/3626772.3657977

[114] Jérémie Rappaz, Julian McAuley, and Karl Aberer. 2021. Recommendation on live-streaming platforms: Dynamic availability and repeat consumption. In *Proceedings of the 15th ACM Conference on Recommender Systems (RecSys'21)*. Association for Computing Machinery, New York, NY, USA, 390–399. DOI: https://doi.org/10.1145/3460231.3474267

[115] Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgilio A F Almeida, and Wagner Meira. 2020. Auditing radicalization pathways on YouTube. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT*'20)*. Association for Computing Machinery, New York, NY, USA, 131–141. DOI: https://doi.org/10.1145/3351095.3372879

[116] Ethan Z Rong, Mo Morgana Zhou, Zhicong Lu, and Mingming Fan. 2022. "It feels like being locked in a cage": Understanding blind or low vision streamers' perceptions of content curation algorithms. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference (DIS'22)*. Association for Computing Machinery, New York, NY, USA, 571–585. DOI: https://doi.org/10.1145/3532106.3533514

[117] Marcelo Luis Barbosa dos Santos. 2022. The "so-called" UGC: An updated definition of user-generated content in the age of social media. *Online Information Review* 46, 1 (2022), 95–113.

[118] Javier Sanz-Cruzado and Pablo Castells. 2018. Enhancing structural diversity in social networks by recommending weak ties. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys'18)*. Association for Computing Machinery, New York, NY, USA, 233–241. DOI: https://doi.org/10.1145/3240323.3240371

[119] Giuseppe Scavo, Zied Ben Houidi, Stefano Traverso, Renata Teixeira, and Marco Mellia. 2017. WeBrowse: Leveraging user clicks for content discovery in communities of a place. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (12 2017). DOI: https://doi.org/10.1145/3134728

[120] Guy Shani and Asela Gunawardana. 2009. *Evaluating Recommender Systems*. Technical Report MSR-TR-2009-159. Retrieved from https://www.microsoft.com/en-us/research/publication/evaluating-recommender-systems/

[121] Thiago Silveira, Min Zhang, Xiao Lin, Yiqun Liu, and Shaoping Ma. 2019. How good your recommender system is? A survey on evaluations in recommendation. *International Journal of Machine Learning and Cybernetics* 10, 5 (2019), 813–831. DOI: https://doi.org/10.1007/s13042-017-0762-9

[122] Ellen Simpson and Bryan Semaan. 2021. For You, or For "You"? Everyday LGBTQ+ Encounters with TikTok. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (1 2021). DOI: https://doi.org/10.1145/3432951

[123] Christian Stöcker and Mike Preuss. 2020. Riding the wave of misclassification: How we end up with extreme YouTube content, Gabriele Meiselwitz (Ed.). Springer International Publishing, Cham, 359–375.

[124] Sajedul Talukder and Bogdan Carbunar. 2020. A study of friend abuse perception in Facebook. *ACM Transactions on Social Computing* 3, 4 (9 2020). DOI: https://doi.org/10.1145/3408040

[125] Chien-Lin Tang, Jingxian Liao, Hao-Chuan Wang, Ching-Ying Sung, and Wen-Chieh Lin. 2021. ConceptGuide: Supporting online video learning with concept map-based recommendation of learning path. In *Proceedings of the Web Conference 2021 (WWW'21)*. Association for Computing Machinery, New York, NY, USA, 2757–2768. DOI: https://doi.org/10.1145/3442381.3449808

[126] Maria Taramigkou, Dimitris Apostolou, and Gregoris Mentzas. 2017. Supporting creativity through the interactive exploratory search paradigm. *International Journal of Human-Computer Interaction* 33, 2 (2 2017), 94–114. DOI: https://doi.org/10.1080/10447318.2016.1220104

[127] Serena Tardelli, Marco Avvenuti, Maurizio Tesconi, and Stefano Cresci. 2020. Characterizing social bots spreading financial disinformation, Gabriele Meiselwitz (Ed.). Springer International Publishing, Cham, 376–392.

[128] Matus Tomlein, Branislav Pecher, Jakub Simko, Ivan Srba, Robert Moro, Elena Stefancova, Michal Kompan, Andrea Hrckova, Juraj Podrouzek, and Maria Bielikova. 2021. An audit of misinformation filter bubbles on YouTube: Bubble bursting and recent behavior changes. In *Proceedings of the 15th ACM Conference on Recommender Systems (RecSys'21)*. Association for Computing Machinery, New York, NY, USA, 1–11. DOI: https://doi.org/10.1145/3460231.3474241

[129] Kristen Vaccaro, Christian Sandvig, and Karrie Karahalios. 2020. "At the end of the day facebook does what itwants": How users experience contesting algorithmic content moderation. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (10 2020). DOI: https://doi.org/10.1145/3415238

[130] Kristen Vaccaro, Ziang Xiao, Kevin Hamilton, and Karrie Karahalios. 2021. Contestability for content moderation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (10 2021). DOI: https://doi.org/10.1145/3476059

[131] Mojtaba Vaismoradi, Jacqueline Jones, Hannele Turunen, and Sherrill Snelgrove. 2016. Theme development in qualitative content analysis and thematic analysis. (2016).

[132] Sebastián Valenzuela, Martina Piña, and Josefina Ramírez. 2017. Behavioral effects of framing on social media users: How conflict, economic, human interest, and morality frames drive news sharing. *Journal of Communication* 67, 5 (10 2017), 803–826. DOI: https://doi.org/10.1111/jcom.12325

[133] Shaofeng Wang, José Paulo Esperança, and Qiao Wu. 2023. Effects of live streaming proneness, engagement and intelligent recommendation on users' purchase intention in short video community: Take TikTok (douyin) online courses as an example. *International Journal of Human-Computer Interaction* 39, 15 (9 2023), 3071–3083. DOI: https://doi.org/10.1080/10447318.2022.2091653

[134] X Wang, L Sun, Z Wang, and D Meng. 2012. Group recommendation using external followee for social TV. In *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo*. 37–42. DOI: https://doi.org/10.1109/ICME.2012.122

[135] Helena Webb, Pete Burnap, Rob Procter, Omer Rana, Bernd Carsten Stahl, Matthew Williams, William Housley, Adam Edwards, and Marina Jirotka. 2016. Digital wildfires: Propagation, verification, regulation, and responsible innovation. *ACM Transactions on Information Systems* 34, 3 (4 2016). DOI: https://doi.org/10.1145/2893478

[136] Janith Weerasinghe, Bailey Flanigan, Aviel Stein, Damon McCoy, and Rachel Greenstadt. 2020. The pod people: Understanding manipulation of social media popularity via reciprocity abuse. In *Proceedings of the Web Conference 2020 (WWW'20)*. Association for Computing Machinery, New York, NY, USA, 1874–1884. DOI: https://doi.org/10.1145/3366423.3380256

[137] Laura Weidinger, John Mellor, Maribeth Rauh, Conor Griffin, Jonathan Uesato, Po-Sen Huang, Myra Cheng, Mia Glaese, Borja Balle, and Atoosa Kasirzadeh. 2021. Ethical and social risks of harm from language models. arXiv:2112.04359. Retrieved from https://arxiv.org/abs/2112.04359. (2021).

[138] Rachel Wood. 2020. "What I'm not gonna buy": Algorithmic culture jamming and anti-consumer politics on YouTube. *New Media & Society* 23, 9 (7 2020), 2754–2772. DOI: https://doi.org/10.1177/1461444820939446

[139] Eva Yiwei Wu, Emily Pedersen, and Niloufar Salehi. 2019. Agent, gatekeeper, drug dealer: How content creators craft algorithmic personas. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019). DOI: https://doi.org/10.1145/3359321

[140] Ruobing Xie, Yanlei Liu, Shaoliang Zhang, Rui Wang, Feng Xia, and Leyu Lin. 2021. Personalized approximate pareto-efficient recommendation. In *Proceedings of the Web Conference 2021 (WWW'21)*. Association for Computing Machinery, New York, NY, USA, 3839–3849. DOI: https://doi.org/10.1145/3442381.3450039

[141] Ming Yan, Jitao Sang, Tao Mei, and Changsheng Xu. 2013. Friend transfer: Cold-start friend recommendation with cross-platform transfer learning of social knowledge. In *Proceedings of the 2013 IEEE International Conference on Multimedia and Expo (ICME)*. 1–6. DOI: https://doi.org/10.1109/ICME.2013.6607510

[142] Ming Yan, Jitao Sang, and Changsheng Xu. 2015. Unified YouTube video recommendation via cross-network collaboration. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (ICMR'15)*. Association for Computing Machinery, New York, NY, USA, 19–26. DOI: https://doi.org/10.1145/2671188.2749344

[143] Xing Yi, Liangjie Hong, Erheng Zhong, Nanthan Nan Liu, and Suju Rajan. 2014. Beyond clicks: Dwell time for personalization. In *Proceedings of the 8th ACM Conference on Recommender Systems (RecSys'14)*. Association for Computing Machinery, New York, NY, USA, 113–120. DOI: https://doi.org/10.1145/2645710.2645724

[144] Brita Ytre-Arne and Hallvard Moe. 2020. Folk theories of algorithms: Understanding digital irritation. *Media, Culture & Society* 43, 5 (12 2020), 807–824. DOI: https://doi.org/10.1177/0163443720972314

[145] Haizi Yu, Biplab Deka, Jerry O Talton, and Ranjitha Kumar. 2016. Accounting for taste: Ranking curators and content in social networks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'16)*. Association for Computing Machinery, New York, NY, USA, 2383–2389. DOI: https://doi.org/10.1145/2858036.2858219

[146] Eva Zangerle and Christine Bauer. 2022. Evaluating recommender systems: Survey and framework. *ACM Computing Surveys* 55, 8 (12 2022). DOI: https://doi.org/10.1145/3556536

[147] Huimin Zeng, Zhankui He, Zhenrui Yue, Julian McAuley, and Dong Wang. 2024. Fair sequential recommendation without user demographics. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'24)*. Association for Computing Machinery, New York, NY, USA, 395–404. DOI: https://doi.org/10.1145/3626772.3657703

[148] Yan Zhao, Shoujin Wang, Yan Wang, Hongwei Liu, and Weizhe Zhang. 2020. Double-wing mixture of experts for streaming recommendations, Zhisheng Huang, Wouter Beek, Hua Wang, Rui Zhou, and Yanchun Zhang (Eds.). Springer International Publishing, Cham, 269–284.

[149] Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, and Ed Chi. 2019. Recommending what video to watch next: A multitask ranking system. In *Proceedings of the 13th ACM Conference on Recommender Systems (RecSys'19)*. Association for Computing Machinery, New York, NY, USA, 43–51. DOI: https://doi.org/10.1145/3298689.3346997

[150] Ruiqi Zheng, Liang Qu, Tong Chen, Kai Zheng, Yuhui Shi, and Hongzhi Yin. 2024. Poisoning decentralized collaborative recommender system and its countermeasures. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'24)*. Association for Computing Machinery, New York, NY, USA, 1712–1721. DOI: https://doi.org/10.1145/3626772.3657814

## A Conferences and Journals Included in the Article Search

Table 6. A List fo Conferences and Journals Included in the Article Search

| **Conference** | **Acronym** | **Publisher** |
|---|---|---|
| ACM CHI Conference on Human Factors in Computing Systems | CHI | ACM |
| ACM Conference on Computer-Supported Cooperative Work and Social Computing | CSCW | ACM |
| ACM on HCI | PACM | ACM |
| ACM Symposium on User Interface Software and Technology | UIST | ACM |
| International Conference on Intelligent User Interfaces | IUI | ACM |
| ACM Conference on Pervasive and Ubiquitous Computing | UbiComp | ACM |
| ACM International Conference on Supporting Group Work | GROUP | ACM |
| Conference on Computing and Sustainable Societies | COMPASS | ACM |
| Conference on Recommender Systems | RecSys | ACM |
| International Conference on Multimodal Interaction | ICMI | ACM |
| Conference on Designing Interactive Systems | DIS | ACM |
| HCI Theory and Applications | HUCAPP | ACM |
| Human Information Interaction and Retrieval | CHIIR | ACM |
| Conference on Hypertext and Social Media | HT | ACM |
| ACM International Conference on Interactive Media Experiences | IMX | ACM |
| HCI International | HCII | Springer |
| IFIP Conference on HCI | INTERACT | Springer |
| Fairness, Accountability, and Transparency | FAccT | ACM |
| Transactions on Interactive Intelligent Systems | TiiS | ACM |
| International Journal of HCI | TJHCI | Taylor and Francis |
| ACM Transactions on Computer-Human Interaction | TOCHI | ACM |
| Universal Access in the Information Society | UAIS | Springer |
| Transactions on Social Computing | TSC | ACM |
| New Media and Society | NMS | Sage Journals |
| International Journal of Human-Computer Studies | IJHCI | Elsevier |

Received 1 December 2024; revised 13 September 2025; accepted 22 September 2025